New Measures for Witnesses: Are They Performing as Anticipated?

Louisa Cliff

Department of Psychology

University of Glasgow

Submitted for the Degree of Ph D. to the Higher Degree Committee of the College of Science and Engineering, University of Glasgow.

April 2011

Abstract

The route of the witness through the criminal justice system, from identification to testimony, has always been fraught with difficulty. There are innumerable sources of influence, intimidation and manipulation which can affect a witness's performance at any point on this pathway. In recent years, and in an effort to resolve several of the issues surrounding witness performance, the British Government have introduced a number of legislative measures to improve performance at identity parades and enhance the court experience of vulnerable witnesses and thereby improve the quality of their testimony. While it seems that these measures have been generally well received, there remain some valid questions over their implementation and whether they are eliciting the desired effects.

The sequential identity parade system, V.I.P.E.R. was devised primarily as a cost-cutting procedural change by West Yorkshire Police in 1999. However, V.I.P.E.R.'s design was heavily influenced by the large extant literature declaring a sequential superiority effect for target absent parades. Specifically, the data claim that sequential identity parades significantly reduce the rate of false identifications from target absent parades whilst preserving hit rates for target present parades. Yet all of the studies purporting this effect have used a very different procedure to that employed by V.I.P.E.R.; these studies use, and recommend, a strict sequential procedure whilst UK legislation requires V.I.P.E.R. parades to be shown twice. Thus the effects of using a V.I.P.E.R. parade upon rates of identification and misidentifications have yet to be empirically investigated. The results of this study revealed that there was no significant effect of parade presentation mode upon rates of correct identifications for target present parades or the rate of mistaken identifications in target absent parades.

The next step for many witnesses is the progression of their case to trial. The incidence of intimidation and vulnerability of adult witnesses appears to be increasing and so a range of Special Measures, more commonly used with child witnesses, were extended to adult witnesses who meet legislative criteria. However, there again has been no investigation into the effects the use of the Special Measures for adult witnesses has upon jurors 'perceptions

of that witness, the defendant and, if there is a change in perceptions, whether they are substantive enough to alter the trial verdict. The current experiments revealed that the use of a screen does not significantly alter jurors' perceptions and had no effect on conviction rates. The use of live CCTV links led the witness to be perceived as less credible but again this did not affect the verdict. It is the use of pre-recorded video evidence which raises the most concern; witnesses testifying in this manner are perceived as significantly less credible and believable which in turn has a significant effect upon conviction rates. However, when tested in a deliberating jury group situation, all the Special Measures were associated with the witness being perceived as more credible compared to the control condition, although these differences were not significant, and there was no significant effect upon the rate of guilty verdicts across all conditions.

Overall, it appears that the implementation of both of these legislative steps have had varying degrees of success. While the V.I.P.E.R. identity parade procedure has reduced running costs it appears that there are no significant differences between V.I.P.E.R. and simultaneous procedures for either target present or target absent identity parades. Conversely, it appears that the introduction of the Special Measures for vulnerable adult witnesses has been largely well received and tentatively successful in reducing witness anxiety. The data suggest that there are no significant negative effects associated with using these Special Measures, certainly under the more ecologically valid condition of the deliberating jury group. Nevertheless, a series of policy recommendations have been made for both identity parade and testimony procedures with the aim of further enhancing the witness's performance and experience through the criminal justice system.

Acknowledgments

I would like to thank Professor Mike Burton for his support, guidance and encouragement throughout the course of these investigations. I would also like to thank Judith McKluskey and Joe Currie for their invaluable assistance in recording the mock trial videos used in Chapter 3. Further thanks are due to Dr Markus Bindemann, Dr Rob Jenkins and Dr David White for their assistance in staging the crime video used in Chapter 2.

Thanks must also go to my parents, Andrew and Louise, for their unending support and encouragement; I certainly would never have reached this stage without you both.

Finally, to Joe; for putting up with many months of stress and anxiety, for being patient and supportive and for making me laugh: thank you, I'm eternally grateful.

Publications

The results of Experiments 1(a)-(c) were presented at The British Criminological Annual Conference 2008, Huddersfield, United Kingdom.

The results of Experiments 1(a)-(c) and Experiment 2 were presented at the 19th Conference of The European Association of Psychology and Law, Sorrento, Italy.

The results of the experiments on identity parade procedures were presented to The British Psychological Society Division of Forensic Psychology Annual Conference 2009.

This work was conducted in fulfilment of a '1+3' Studentship from the Economic and Social Research Council; award number PTA-031-2005-00334.

Declaration

This thesis has been composed by the undersigned. It has not been accepted in any previous application for a degree. The work, of which this thesis is a record, has been completed by myself, unless otherwise indicated in the text. I further state that no part of this thesis has already been, or is concurrently, submitted for any such degree or qualification at any other university.

……

Louisa Cliff

Table of Contents

Chapter 1:
General Introduction

Introduction

The route of the eye-witness through the criminal justice system is often a frightening and frustrating experience requiring repeated questioning, on-demand attendance at identity parades and a confrontational examination in court. It is therefore not surprising that more than 75% of witnesses express dissatisfaction and frustration with their experience within the justice system (Richards, Morris & Richards, 2008; Slater, 1995; Hoyano, 2001). This discontent appears to stem from two sources; firstly witnesses and support groups frequently describe a sense of abandonment and lack of information from judicial organisations (Richards, Morris & Richards, 2008; Hamlyn, Phelps, Turtle & Sattar, 2004) and, secondly, frustration at the apparent lack of procedural considerations when attending identity parades and testifying in court (Richards, Morris & Richards, 2008; Slater, 1995; Steblay, Dysart, Fulero & Lindsay, 2001; Home Office Research Study 283, June 2004; Wells, 2001; Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003; Wheatcroft, Wagstaff & Kebbell, 2004).

With regards to the first issue there is certainly evidence of a breach between witnesses and the justice system. It has been recognised by the government, police and judges that all too often witnesses report a crime and then hear nothing further until a citation to attend court is delivered (Richards, Morris & Richards, 2008). This feeling of disregard often extends after conclusion of the trial whereby some witnesses are not even informed of the verdict and, if relevant, sentencing. In recent years however, several agencies have been formed in an effort to address such issues, improving the witness experience whilst minimising the prevalent sense of neglect. Recent review of the witness experience in British courts suggest that, while there have been a range of procedural improvements, there remains an acute lack of contact between the police and judicial services and witnesses (Richards, Morris & Richards, 2008; Burton, Evans & Sanders, 2006; Hamlyn *et al.*, 2004).

Historically the emphasis of all judicial proceedings, from investigation to trial, has been to ensure a fair, unbiased trial for the accused. While such a statement holds today there is also an assumption of duty to protect the physical and emotional well-being of victims and

witnesses throughout the judicial system. Previously there was always a possibility of the witness encountering the accused at the identity parade or while attending court; situations which would understandably be intimidating, possibly terrifying, and certainly detrimental to the witness's performance in these tasks (Slater, 1995; Steblay *et al.,* 2001; Richards, Morris & Richards, 2008; Regan & Baker, 1998; Golding, Fryman, Marsil & Yozwiak, 2003; Goodman, Tobey, Batterman-Faunce, Orcutt, Thomas, Shapiro & Sachsenmaier, 1998; Birch, 2000; Hoyano, 2001). Recent legislation however has eliminated the possibility of a chance meeting of witness and accused by revising both the format of the identity parade system and by providing access to protective 'Special Measures' during testimony for the most severe crimes and most vulnerable of witnesses (The Police and Criminal Evidence Act 1984, Code D 2008; The Vulnerable Witnesses (Scotland) Act, 2004, The Youth Justice and Criminal Evidence Act 1999).

If we address first the change in identity parade procedures, the benefits to the witness are immediately evident; there is no longer *any* possibility of an encounter between witness and accused and the parade can be administered outside the police station in a more comfortable location for the witness (Slater, 1995; Steblay *et al.,* 2001). Such a step would undoubtedly improve the witness's performance compared to the established, more traditional parade format whilst also contributing to a reduction in police expenditure and resources (Slater, 1995). The switch to a sequential format has long been advocated by researchers (Lindsay & Wells, 1985; McQuiston-Surrett, Malpass & Tredoux, 2006; Gronlund, 2004; Kassin, Tubb, Hosch & Memon, 2001; Steblay *et al.,* 2001; Wells, Small, Penrod, Malpass, Fulero & Brimacombe, 1998) yet, as we will later discuss, there is continuing debate regarding the specific conditions under which the new V.I.P.E.R. parades (*Video Identity Parade, Electronic Recording*) are superior to the more traditional, simultaneous parade format (Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Wells, 2001).

Equally, steps have been taken to improve the witness experience at court. Appearing in court before the accused and a large group of strangers to deliver your testimony is an

incredibly disquieting experience and one which often exerts an often debilitating effect upon witness demeanour and performance (Golding *et al.,* 2003; Regan & Baker, 1998). The introduction of The Vulnerable Witnesses (Scotland) Act 2004 and the 1999 Youth Justice and Criminal Evidence Act in England and Wales has extended the inclusion criteria for witnesses who may benefit from 'Special Measures' during testimony (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003). Such protective measures in the court are by no means a new initiative however; previously their use had been reserved for child witnesses or adult witnesses suffering from learning difficulties (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003). While these witnesses remain assured of their right to protective measures, the Acts have recognised the growing problem of witness intimidation (Criminal Law Review Editorial, 2004; Richards, Morris & Richards, 2008) and thus permit vulnerable and/or intimidated adult witnesses to apply for special consideration under certain conditions (The Vulnerable Witnesses (Scotland) Act, 2004; Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003). However, the provisions now permitted for adult witnesses were implemented without empirically examining the effects that these 'Special Measures' exert upon jurors' perceptions of adult witnesses and defendant guilt.

It is apparent then that the changes in identity parade and trial procedures, whilst beneficial from the witness's viewpoint, has yet to be scientifically tested to fully examine the consequences these legislative steps have had upon the success of identity parades, the perceptions of jurors and the efficacy of the justice system as a whole. It is the aim of this thesis to investigate the effects and determine the wider implications elicited by the change in identity parade format and the use of 'Special Measures' by adult vulnerable witnesses.

Identity Parades

The image of an identity parade is an indelible component of our ideal criminal justice system and is often believed to be an integral step towards solving a crime. Identity parades are frequently assumed by the public to be fair and free from biasing influences yet

extensive research from the previous 30 years has shown such assumptions to be false (Lindsay & Wells, 1985; Valentine & Heaton, 1999; Wells, 1984, 1993; Valentine, 2006; Valentine, Harris, Piera & Darling, 2003; Malpass & Lindsay, 1999; Valentine & Heaton, 1999). Identity parades, while valuable to the investigative process, rely upon the recall and recognition abilities of the eyewitness (Wells, 1984; Bruce, Henderson, Newman & Burton, 2001; Deffenbacher, Bornstein, Penrod & McGorty, 2008; Megreya & Burton, 2008). There is a vast range of factors which may influence the recognition and identification skills of the eyewitness but these are typically divided into two broad groups of *system* variables, which may be controlled, and *estimator* variables, which are typically beyond the control of the justice system (Lindsay & Pozzulo, 1999; Memon & Gabbert, 2003; Wells, 1978; Wells & Olson, 2003).

Estimator variables are factors which are beyond the control of the criminal justice system. Such influences include environmental factors when the crime was committed; light quality, distance from suspect but also extends to several more complex psychological effects such as the other race effect (Pezdek, Blandon-Gitlen & Moore, 2003, Sporer, 2001, Meissner & Brigham, 2001), the difficulty associated with unfamiliar face matching (Megreya & Burton, 2006, Megreya & Burton, 2008, Kemp, Towell & Pike, 1997, Valentine, Darling & Memon, 2007, Newell, Chiroro & Valentine, 1999), the presence of a weapon (Steblay, 1992, MacLin, MacLin & Malpass, 2001, Maass & Kohnken, 1989), the rapid decay of memory (Deffenbacher *et al.,* 2008, MacLin *et al.,* 2001) and the level of personal stress experienced by the witness throughout the incident (Valentine & Mesout, 2009, Morgan, Hazlett, Doran, Garrett, Hoyt, Thomas, Baranoski & Southwick, 2004, Deffenbacher, Bornstein, Penrod & McGorty, 2004).

The probability of obtaining a successful, accurate and reliable identification from the eye-witness declines significantly with each of the effects mentioned above. At the root of the matter is the inherent difficulty in recognising and matching unfamiliar faces. There is substantial evidence which supports this effect (Megreya & Burton, 2008, Kemp *et al.,* 1997, Valentine *et al.,* 2007, Megreya & Burton, 2006, Hancock, Bruce & Burton, 2000, Newell,

Chiroro & Valentine, 1999, Bruce, Henderson, Newman & Burton, 2001) and witnessing a crime is the most common area we would expect to see this effect. When a crime is witnessed it is relatively unlikely that the witness will have prior knowledge of the perpetrator; criminal acts are planned to minimise witnesses but certainly to avoid locations where the perpetrator is known (unless a disguise is employed). Therefore when the eye-witness is placed before the parade they are being asked to recognise an unfamiliar person who they have only seen once previously for perhaps only a short time (Deffenbacher *et al.,* 2008, MacLin *et al.,* 2001, Memon, Hope & Bull, 2003, Lindsay & Pozzulo, 1999, Read, Vokey & Hammersley, 1990) and under intense pressure (Morgan *et al.,* 2004, Valentine & Mesout, 2009, Deffenbacher *et al.,* 2004).

There is a compelling body of evidence to support the poor performance of witnesses when asked to recognise an unfamiliar face. Kemp, Towell and Pike (1997) asked experienced cashiers to match shoppers currently before them to their credit card photographs and reported a significantly high rate of false identifications even though conditions had been optimised for successful identification. This is in stark contrast to the apparent ease with which participants can identify familiar individuals, generating significant identification rates even from poor quality CCTV images which have degraded (Hancock, Bruce & Burton, 2000, Bruce *et al.,* 2001, Burton, Wilson, Cowan & Bruce, 1999). It has been suggested that the poor performance regarding unfamiliar face matching and recognition stems from a difficulty in encoding the unfamiliar face whereby the unfamiliar face is not perceived as a face but rather as a simple pattern (Hancock *et al.,* 2000, Megreya & Burton, 2006, 2008). Due to the inability to perceive the unfamiliar image as a face the specialised encoding strategies are not engaged and therefore lead to poor matching and recognition abilities later (Megreya & Burton, 2006, Wells, 1978).

There is also a strong negative effect upon identification when the perpetrator is of a different race to the eye-witness (Wells & Olson, 2001, Smith, Lindsay, Pryke & Dysart, 2001, Goldstein & Chance, 1979, Brigham & Barkowitz, 1978, Pezdek *et al.,* 2003, Sporer, 2001, Meissner & Brigham, 2001). This effect is frequently associated most strongly with

those racial groups with which the eye-witness has the least contact (Meissner & Brigham, 2001, Wright, Boyd & Tredoux, 2003) and is often typified by the claim of 'They all look alike to me' (Goldstein & Chance, 1979, Brigham & Barkowitz, 1978, Malpass, 1981, Ferguson, Rhodes, Lee, & Sriram , 2001, Johnson & Fredrickson, 2005). The Other-Race Effect or Cross-Race Bias is a robust finding based upon significant research from across the world (Sporer, 2001, Meissner & Brigham, 2001, Johnson & Fredrickson, 2005, Wright *et al.,* 2003) and this deficit in recognition has been documented to affect eye-witnesses of all races (Meissner & Brigham, 2001, Wright *et al.,* 2003). Thus same-race faces would be expected to persist in the eye-witness's memory for a longer period than a different-race face (Wells & Olson, 2003, Smith *et al.,* 2001, Meissner & Brigham, 2001, Wright *et al.,* 2003, Ferguson *et al.,* 2001, Pezdek *et al.,* 2003).

The presence of a weapon (Steblay, 1992, MacLin, MacLin & Malpass, 2001, Maass & Kohnken, 1989, Wells, 1978, Kramer, Buckhout & Eugenio, 1990) or additional unfamiliar faces further debilitates the eyewitness's later recognition skills. In their 2006 study, Megreya & Burton examined the effects of viewing more than one target face upon later identification of a single target face from a photo array. Their findings show clearly that the rate of successful identification under these conditions was just 34% in target present parades and suggests that the presence of two unfamiliar faces simultaneously is associated with poor perceptual performance with one of the two faces always receiving more attention than the other. Such a division of attention leads to poor encoding of the two faces and thus negative affects the eye-witness's recognition ability. When the study was repeated using a sequential, rather than simultaneous presentation, of the two unfamiliar target faces performance was significantly worse than that observed from the simultaneous experiment; a finding, which we will discuss in more depth later, corroborates data from sequential identity parade research (Lindsay & Pozzulo, 1999). Megreya & Burton's (2006) findings reveal a significant recognition disadvantage when presented with two faces however the cause of this effect; whether it is a result of memory failure or poor encoding, remains elusive.

The ‘weapon focus’ effect has been widely documented in forensic research (Steblay, 1992, MacLin, MacLin & Malpass, 2001, Maass & Kohnken, 1989, Wells, 1978, Kramer *et al.,* 1990) and is linked to the associated ‘stress’ factor which can also affect eye-witness performance at identity parades (Valentine & Mesout, 2009, Morgan *et al.,* 2004, Deffenbacher *et al.,* 2004, Pozzulo, Crescini & Panton, 2008). Firstly, ‘weapon focus’ is the theory that the presence of a weapon, particularly a lethal weapon, draws eye-witness attention away from the perpetrator’s face and towards the weapon resulting in little encoding of the face to aid in later recognition (Yarmey & Jones, 1983, Valentine & Mesout, 2009, Morgan *et al.,* 2004, Deffenbacher *et al.,* 2004, Steblay, 1992). Indeed in their 1983 review, Yarmey & Jones claimed that 88% of researchers believed that the presence of a weapon significantly reduced the likelihood of a positive identification. Steblay’s 1992 meta-analysis of nineteen studies of the weapon effect revealed a highly significant reduction in identification when a weapon was used in the crime. Steblay’s review asserts that this effect is prompted by the attentional bias drawn towards the weapon which leads to a lack of encoding of secondary details including the facial features of the perpetrator, which prompts later poor performance at an identity parade.

The presence of a weapon may also be a contributory factor to a further estimator variable which affects eye-witness recognition skills; stress. The level of emotional stress experienced by the witness during the commissioning of the crime is thought to be a considerable influencing factor which debilitates identification (Morgan *et al.,* 2004, Valentine & Mesout, 2009, Deffenbacher *et al.,* 2004, Pozzulo *et al.,* 2008). The consensus of the current literature is that recognition may be severely reduced when the eye-witness encounters an unfamiliar face during a period of high stress, but that this is very much dependent upon the exact situation encountered. This theory has been tested extensively; Valentine & Mesout’s 2009 study took place at the London Dungeon and attempted to link their findings with physiological responses. Participants were asked to wear a heart monitor whilst travelling through the London Dungeon and at some point they encountered an actor who blocked their path. The results show a clear link between physical anxiety and recognition with those participants who experienced high anxiety performed significantly

poorer on the identification task while those who were less anxious were more likely to make a correct identification.

Stress and arousal has been a key factor in a number of research studies investigating eyewitness fidelity yet there is still no clear picture of its exact pattern of influence. In an attempt to clarify the effect size and direction of stress on eyewitness memory, Deffenbacher *et al.* (2004) conducted a meta-analysis of the literature. In his review of 34 published studies, Deffenbacher determined that high levels of stress during a crime incident negatively impacted upon both eyewitness identification accuracy and the accuracy of crime-related details. Specifically, the rates of correct identifications of the suspect from a culprit present parade was significantly reduced, leading Deffenbacher and colleagues to propose that high stress levels was linked with increased memory degradation which in turn leads to decreased ability to make a correct identification of the culprit. A related finding of the meta-analysis was that studies conducted in the field, in a more ecologically valid situation, were associated with a greater level of perceived threat compared to those studies conducted in a laboratory. Deffenbacher *et al.* claim that the size of the stress effect from laboratory based studies may be providing a significant underestimation of the incapacitating effects of heightened stress due to their inability to achieve stress levels observed in field studies.

Clearly studies investigating the effects that estimator variables can exert upon eyewitness recognition abilities have produced valuable data for the forensic researcher, yet the applicability and relevance of such studies may be somewhat limited. While the above mentioned variables are only a small selection of influencing factors it is apparent that estimator variables are impossible to control in a real life situation (Wells, 1978, Wells & Olson, 2003, Valentine 2006). There have been suggestions in the literature that the research examining estimator variables could be used to formulate a quantitative estimate of the accuracy of any one eyewitness which can be later referenced in court or perhaps be employed as a cautionary guide which a psychologist would deliver to jurors before trial commences (Wells, 1978). Yet such ideas have been widely discredited because there is

little empirical evidence which suggests jurors over-believe eye-witnesses (Penrod & Cutler, 1995, Wells & Seelau, 1995, Wells & Olson, 2003). In fact there is significant evidence to suggest that jurors are not capable of recognising the limitations of eyewitnesses and their testimony (Wells & Olson, 2003, Wells, 1978, Quas, Thomson & Clarke-Stewart, 2005, McAuliff & Kovera, 2007; Semmler, Brewer & Douglass, in press). McAuliff & Kovera (2007) examined the differences between expert and lay knowledge of witness suggestibility research. They asked 58 expert psychologists, 157 jurors and 220 undergraduate students to estimate the effects of misleading information across a wide range of conditions and situations. Their results demonstrated that knowledge of the adverse effects of delays, the positive effects of pre-misinformation warnings and age-related differences was similar across all participants. The lay participants rated themselves as being unfamiliar with much of witness suggestibility research and felt that expert testimony in complex cases would be of benefit to jurors. Such findings are most succinctly illustrated by the statistics of The Innocence Project. The Innocence Project is an organisation in the United States dedicated to "exonerating wrongfully convicted individuals through DNA testing" (http://www.innocenceproject.org/know/) and, to date, has successfully exonerated 268 individuals wrongfully convicted on the basis of flawed eyewitness testimony. Of these, 17 had been sentenced to death before their innocence was proved and the average time spent in prison by the exonerees before release is 13 years. This is perhaps the most graphic reminder of the fallibility of eyewitness testimony and its human cost.

Conversely, system variables are the factors which can and should be controlled by the criminal justice system. Within this group are factors such as the type of identification procedure employed: photo-spreads, simultaneous or sequential parades (Bromby, 2002; Valentine, Darling & Memon, 2007; McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Bruce *et al.,* 2001; Wells, 2001; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999; Valentine, 2006; Wells, 1984; Lindsay & Wells, 1985; Busey & Loftus, 2007; Tredoux & Chiroro, 2006; Malpass & Lindsay, 1999; Valentine & Heaton, 1999; Ebbesen & Flowe, 2002; Kemp, Pike & Brace, 2001); the selection of parade foils (Lindsay & Pozzulo, 1999; Slater, 1995; Steblay *et al.,* 2001; Wells & Olson, 2003; Clark & Tunnicliff, 2001; Luus & Wells, 1991);

administration of the parade (Kemp *et al.,* 2001; Tredoux & Chiroro, 2006; Levi & Lindsay, 2001; Steblay *et al.,* 2001; Wells *et al.,* 1998; McQuiston-Surrett *et al.,* 2006); instructions given to the witness (Valentine *et al.,* 2007; Tredoux & Chiroro, 2006; Steblay, 1997; Malpass & Devine, 1981; McQuiston-Surrett *et al.,* 2006; Wells *et al.,* 1998) and post-identification feedback (Douglass & McQuiston-Surrett, 2006; Lampinen, Scott, Pratt, Leding & Arnal, 2007; Tredoux & Chiroro, 2006; Charman & Wells, 2008).

The selection of suitable foils is a significant factor which should be strictly controlled by the organising police force in order to ensure fairness to the suspect (Lindsay & Pozzulo, 1999, Slater, 1995, Steblay *et al.,* 2001, Wells & Olson, 2003, Clark & Tunnicliff, 2001, Luus & Wells, 1991). Every effort should be made to locate and recruit similar individuals to complete the parade, which appears to be a relatively simple task, yet Slater (1995) reported that eye-witnesses felt that less than 50% of the foils used were similar to the target individual. The selection of appropriate foils is in actuality a very difficult task for police forces. If a traditional, simultaneous parade is to be used officers are constrained to selecting individuals from the local area who match a written description of the suspect (Steblay *et al.,* 2001; Slater, 1995). Thus it is not surprising that foils are routinely said to be a poor match for the perpetrator as there are limitations to the availability of acceptable, and willing, foils in a local district. This is a serious factor affecting eyewitness identification; if the suspect is described as very tall and the additional parade foils are all shorter than the suspect this would have a biasing effect upon the witness's identification.

Recent changes in identity parade procedures have addressed such issues with the introduction of the V.I.P.E.R. identity parade procedure. As parades are no longer conducted live there is no need to search for locally available suitable foils; there is now a national database of 12,000 images which can offer a wide selection of the most appropriate foils. In an additional benefit, this provision of 'known innocent' foils from a national database significantly minimises the chance of a false identification progressing into a wrongful conviction.

The choice of which identity parade procedure is employed is a further system variable which is recommended for judiciary control (Wells, 2001; Wells & Olson, 2003; Lindsay & Pozzulo, 1999). Recently the sequential V.I.P.E.R. (**V**ideo **I**dentity **P**arade, **E**lectronic **R**ecording) has been implemented as first-choice procedure for identity parades in the UK. Sequential presentation of identity parades is advocated widely in the literature with over 80% of researchers supporting the format (Kassin, Tubb, Hosch & Memon, 2001; McQuiston-Surrett *et al.,* 2006; Valentine *et al.,* 2007). While simultaneous identity parades are recognised to elicit a significant rate of correct identifications for target present parades, they are also associated with a significant rate of false identifications for target absent identity parades (Steblay *et al.,* 2001; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Memon & Gabbert, 2003; Wells, 2008, 1993; Valentine, 2006; Pike *et al.,* 2000; Levi, 1998; Kassin *et al.,* 2001; Gronlund, 2004). Conversely, there is a substantial body of evidence which claims that sequential presentation of an identity parade significantly reduces the false positive rate, and therefore improves the rate of correct rejections, for target absent parades whilst maintaining the hit rate for target present parades (Steblay *et al.,* 2001; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Wells, 2008, 1993; Valentine, 2006; Pike *et al.,* 2000; Levi, 1998; Kassin *et al.,* 2001; Gronlund, 2004). Thus it is of apparent benefit, without any caveats, that a sequential format be implemented unwaveringly throughout the criminal justice system (Steblay *et al.,* 2001; Valentine, 2006; Wells, 2001; Wells *et al.,* 1998).

However, such a recommendation would be enhanced by the effective control of parade administration over two levels; blind administration (Kemp *et al.,* 2001; Tredoux & Chiroro, 2006; Levi & Lindsay, 2001; Steblay *et al.,* 2001; Wells *et al.,* 1998; McQuiston-Surrett *et al.,* 2006) and witness instructions (Valentine *et al.,* 2007; Tredoux & Chiroro, 2006; Steblay, 1997; Malpass & Devine, 1981; McQuiston-Surrett *et al.,* 2006; Wells *et al.,* 1998). It has previously been common that the administration of the identity parade was performed by a member of the investigating team. Clearly there is a conflict of interest here whereby the administrating officer may have a preconceived suspicion of guilt going into the parade and may therefore, consciously or subconsciously, influence the decision of the witness (Wells *et al.,* 1998; McQuiston-Surrett *et al.,* 2006; Steblay, 1997). It has been extensively suggested

in research over the last thirty years that administration of all identity parades, regardless of the format employed, should be double blind; that the administrating officer is not aware of the presence of any suspect at the time of the parade (Levi & Lindsay, 1999; Wells, 2001; Steblay *et al.,* 2001; Wells, 1993; Lindsay & Wells, 1985).

In recent years there has been a significant degree of controversy over the field testing of double-blind testing in Illinois. A unique opportunity arose in 2002 as the Illinois State Government proposed to issue a report that aimed to determine why the state had such a high number of wrongful convictions and looked to reform the Illinois criminal justice system in the light of the report's findings. Flawed eyewitness testimony was to form a significant part of this report due its well-documented fallibility. Sherri Mecklenburg was approached to conduct a large-scale investigation into the effectiveness of the sequential, double-blind procedure in a year-long pilot study conducted in conjunction with three regional police departments. The results of this study, published in 2006, claimed that these first field tests of the sequential, double-blind procedure were associated with a lower rate of correct identifications of the suspect in target present parades but also with a higher rate of known false identifications (Mecklenburg, 2006). Immediately there was a furore surrounding the conclusions of the report and serious questions were raised by researchers concerning the methodologies employed whilst others proclaimed it a wasted opportunity (Wells, 2006; Steblay, 2006; Schacter *et al.*, 2008; Steblay, 2008; 2010). While the sequential parades had been run double-blind, the simultaneous parades, the control group, had not been (Mecklenburg, 2006; Steblay, 2010; Wells, 2006; Steblay, 2006; Schacter, Dawes, Jacoby, Kahneman, Lempert, Roediger & Rosenthal, 2008). This was not the extent of the problems however; Nancy Steblay obtained the empirical data from one of the testing cities (Evanston) and conducted her own review and analysis of Mecklenburg's (2006) data. Steblay reported some startling methodological differences between the sequential and simultaneous parades. Whilst the sequential format was comprehensively scripted the simultaneously parades were subjected to no efforts of standardisation across participants. The simultaneous parades were not administered double-blind but also were more likely to use witnesses that had already identified the suspect in an earlier identity parade, were more

likely to use witnesses who had a prior relationship with the suspect and the delay between event and recall were much shorter. Steblay (2010) asserts that the Mecklenburg field studies were not an effective strategy to empirically investigate the effectiveness of the double-blind, sequential format; a conclusion shared by many researchers (Wells, 2006; Schacter *et al.*, 2008; Steblay, 2006) including a seven person committee of several prominent researchers (Schacter *et al.*, 2008). Yet it must be noted that the Mecklenburg Report (2006) is not without its support; Ross & Malpass (2008) conducted their own review of the findings of the report and determined that the absence of the 'double-blind' administration in the simultaneous parades is not a debilitating as other researchers claim (Steblay, 2010; Wells, 2006; Steblay, 2006; Schacter *et al.*, 2008) and that confounds exist in laboratory based studies as well as field studies. Instead, Ross & Malpass assert that the Illinois Field Tests are wrongly interpreted as a conclusive means to resolve the Simultaneous/Sequential debate; they recognise that the methodology was flawed and that future field studies must be based upon detailed information from the actual practical use of identity parades and from the case files. The two opposing groups of this continuing controversy regarding double-blind administration is united on one point however; that continuing arguments over past field studies is counter-productive and that the same urgency be applied to developing future research to resolve the issues surrounding double-blind administration of identity parades.

The issue of double-blind identity parades have been resolved with the implementation of the V.I.P.E.R. procedure. The administration of V.I.P.E.R. parades is handled by a team of dedicated parade officers from selection of foils to delivering the parade to witnesses (The Police and Criminal Evidence Act 1984, Code D 2008) and it is specified in law that the administrating officer may not be a member of the investigating team. Equally it is enshrined in law that the administrating officer, in their instructions to the witness, must highlight that the perpetrator may not be present in the parade (The Police and Criminal Evidence Act (1984) Code D, 2008). Such instruction is designed to counter-act the implicit priming said to be associated with attending an identity parade (Wells, 1993); 'They must have caught the

guy if they want me to come to an identity parade' which can encourage the eyewitness towards making an identification even thought the perpetrator may not be present.

The effect of Post-Identification Feedback has been extensively researched and it is a widely accepted theory that providing feedback upon witness performance at an identity parade can significantly interfere with retrospective judgements and confidence (Douglass & McQuiston-Surrett, 2006; Lampinen *et al.,* 2007; Tredoux & Chiroro, 2006; Charman & Wells, 2008; Douglass & Steblay, 2006; Dixon & Memon, 2005). The standard guidelines for best practice at identity parades, as advocated by the research community (Wells *et al.,* 1998; Wells, 2001; Levi & Lindsay, 1999; Steblay *et al.,* 2001; Luus & Wells, 1994) includes obtaining a statement of witness confidence in their identification at the time of the parade. However, it has been expressly shown that even a casual, confirmatory comment from the administrating officer can have an enormous effect upon the eyewitness's later confidence reports but can also affect witnesses' reflections upon their own performance, boosting their apparent 'skill' at recognising the perpetrator (Douglass & McQuiston-Surrett, 2006; Lampinen *et al.,* 2007; Tredoux & Chiroro, 2006; Charman & Wells, 2008; Douglass & Steblay, 2006; Bradfield & Wells, 2005). Thus witnesses who are given positive feedback often claim to have 'immediately' recognised the suspect (Wells & Bradfield, 1998; Bradfield & Wells, 2005; Lampinen *et al.,* 2006; Wells, Olson & Charman, 2002), to have required minimal time to make an identification (Lampinen *et al.,* 2007; Lindsay & Pozzulo, 1999; Douglass & Steblay, 2006) and to have a strong general ability to recognise unfamiliar faces (Douglass & Steblay, 2006). In their 2006 meta-analysis Douglass & Steblay analysed data from 14 studies and reported that participants who received confirmatory feedback rated their confidence in their decision significantly higher than those who received no feedback. In addition, the participants who were given feedback were significantly more likely to claim a special ability to recognise faces, that they had an enduring image of the perpetrator's face and to report their recognition as 'immediate'. Therefore we can see that such an effect would have a serious result if the witness identified an innocent individual and then received positive feedback; potentially causing a miscarriage of justice.

Douglass & Steblay (2006) thus recommend that a fully blind procedure is used in identity parades whereby the administrating officer should have no prior knowledge if the suspect is present in the parade to eliminate any possibility of the witness receiving feedback after an identification has been made; a conclusion supported by many researchers (Wells *et al.,* 1998; Lampinen *et al.,* 2006; Douglass & McQuiston-Surrett, 2006; Wells, 1993; 2001; Levi & Lindsay, 1999; Steblay *et al.,* 2001; Luus & Wells, 1994; Bradfield & Wells, 1998; 2005; Lindsay & Pozzulo, 1999; Wells, Olson & Charman, 2002).

The final system variable that the judicial system is strongly advised to control is the retention interval between witnessing a crime and attending an identity parade (Wells, 1978; Wells & Olson, 2003; Slater, 1995; Steblay *et al.,* 2001). It is proven that memory and recognition begins to decay immediately after exposure ends (Ebbinghaus, 1885; Deffenbacher *et al.,* 2004; Ebbesen & Rienick, 1998; Kemp *et al.,* 2001; Bromby, 2002; Dunning & Stern, 1992; Christianson, 1992). Memory decay is typically Ebbinghausian in nature, forgetting at a startling rate over the first 24 hours and before beginning to level off around 2 days after exposure (Deffenbacher *et al.,* 2004; Ebbinghaus, 1885). Thus there is an obvious pressing need to organise and conduct identity parades as soon as possible after the incident. We can say then that the retention interval associated with traditional simultaneous parades, averaging ten weeks (Valentine, 2006; Valentine, Pickering & Darling, 2003) between encoding and recall, is liable to be associated with particularly poor performance. In contrast, the sequential format has an average retention interval of ten days (Valentine *et al.,* 2003), although can be assembled in hours; which, while still not ideal, would most likely be associated with greater performance at the identity parade. Of course, delay is one system variable which, despite best efforts, cannot always be adequately controlled. In addition to the time delay involved in composing the identity parades, outlined above, the key point at which delay cannot be controlled is the interval between the criminal event and the witness coming forward to the police. Therefore, any identity parade procedure which may reduce further lag between event and retrieval would be greatly beneficial.

It is easy to see then that research into system variables and their effect upon eyewitness identification is of greatest applicability to the criminal justice system (Wells, 1978; Wells & Olson, 2003; Valentine, 2006; Lindsay & Pozzulo, 1999). While estimator variables are very difficult to control, and research into these variables is of perhaps less benefit to practitioners, research into system variables may yet yield some directions to bypass the uncontrollability of estimator variables and thus further improve the identity parade procedures (Wells, 1978; 1984; Wells & Olson, 2003). While many of the system variables have been addressed with the recent introduction of the sequential V.I.P.E.R. parade procedure it is important that we first assess the older, simultaneous identity parade procedure; its benefits and its limitations.

Traditional Identity Parades

The enduring image; eight similar individuals lined up, side-by-side behind mirrored glass, is no longer wholly accurate when called to view an identity parade. This simultaneous format has been employed worldwide certainly for the last century and has only been replaced in the UK within the last ten years despite growing claims for a sequential procedure from the early 1980's (Levi, 1998; Levi & Lindsay, 2001; Lindsay & Wells, 1985; McQuiston-Surrett *et al.,* 2006; Kassin *et al.,* 2001; Steblay *et al.,* 2001; Slater, 1995; Valentine, 2006; Valentine *et al.,* 2007; Wells *et al.,* 1998). So why were police procedures changed after 30 years of research? Traditional parades are acknowledged as an excellent procedure for obtaining an identification but only when the perpetrator is present in the line-up (Lindsay & Wells, 1985; Wells, 1993; Memon & Gabbert, 2003; Steblay *et al.,* 2001; Tredoux & Chiroro, 2006; Lindsay & Pozzulo, 1999; Levi & Lindsay, 2001). This success is claimed to be due to the judgement style simultaneous presentation engenders; a relative judgement style where the eye-witness compares the parade members across each other whilst comparing them individually to their memory of the perpetrator (Lindsay & Wells, 1985; Tredoux & Chiroro, 2006; Levi & Lindsay, 2001; Memon & Gabbert, 2003; McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Valentine *et al.,* 2006; Gronlund, 2004). When the suspect is present therefore, this is a reasonable manner of processing and matching the 'best-fit' face with the

witness's memory. However, this judgement style is claimed to be the source of the serious mis-identification rate that traditional parades elicit when the suspect is not present in the line-up (Valentine, 2006; Steblay *et al.,* 2001; Lindsay & Wells, 1985; McQuiston-Surrett *et al.,* 2006; Memon & Gabbert, 2003; Wells, 2008). By making a relative judgement the eye-witness is identifying an innocent foil, an identification which can lead to severe repercussions for the innocent accused and the criminal justice system. A false positive identification in a traditional, simultaneous parade can frequently lead to prosecution, particularly if the eye-witness is the only evidence source, because there is no system in place where the administrating police officer can be absolutely certain that the recruited foils *are* innocent (Valentine & Heaton, 1999; Valentine, 2006; Steblay *et al.,* 2001; Slater, 1995; The Police and Criminal Evidence Act 1984, Code D 2008) and thus discount a false positive identification.

While the traditional format is superior for target present parades from a policing viewpoint, there are a great many other issues regarding their use; namely the sheer expense and demand on resources they require. As traditional identity parades are carried out in a local police station the pool of available, and appropriate, foils is substantially narrowed. This then requires a significant undertaking to locate and recruit suitable foils every time an identity parade is organised, taking police resources from more serious issues. Not only is this a drain on personnel resources but also financially; each recruited foil is paid £15 for the first hour and £5 for each subsequent hour they are in attendance (West Yorkshire Police).

The costs to the organising police force may be compensated in some fashion by the successful administration of the parade, frequently however this is not the case. Slater (1995) reported that at least 50% of all traditional identity parades were cancelled prior to administration, primarily due to witness concerns over coming face-to-face with the perpetrator (Steblay *et al.,* 2001; Slater, 1995; Valentine, 2006).

The extensive organising process required for traditional identity parades is the primary contributory factor to the often lengthy retention interval between incident and recall. While

this interval is an average of ten weeks it can frequently be longer. As discussed earlier such a prolonged delay is counter-productive to eye-witness performance (Wells, 1978; Wells & Olson, 2003; Slater, 1995; Steblay *et al.,* 2001; Ebbinghaus, 1885; Deffenbacher *et al.,* 2004; Ebbesen & Rienick, 1998; Kemp *et al.,* 2001; Bromby, 2002; Dunning & Stern, 1992; Christianson, 1992; Valentine, 2006; Valentine *et al.,* 2003). This, in combination with the relative judgement style fostered by simultaneous presentation, may serve to hinder the witness's ability to recognise that the perpetrator is not present thus resulting in a mis-identification.

Due to these limitations associated with traditional, live, simultaneous identity parades, West Yorkshire Police devised a sequential, video-based identity parade format, V.I.P.E.R., to improve the witness's experience and performance whilst addressing the general drain upon resources.

Video Identity Parades

Sequential, video identity parades began to be used in the UK in 1999 and within 6 years they were widely used across most British police forces. Legislative measures have placed V.I.P.E.R. identity parades as the preferred method for eyewitness identification, to be used in all conditions unless there is some constraint which makes this impossible (The Police and Criminal Evidence Act 1984, Code D 2008; Valentine *et al.,* 2007).

It is important to note that V.I.P.E.R. is not simply a sequential presentation of static images, photos or even live individuals. V.I.P.E.R. is the sequential presentation of moving video images of a set of similar individuals. It is of course possible to present video clips of faces simultaneously just as it is possible to present live identity parades sequentially. These possible identity parade formats are not permitted under UK law and are not investigated in the current studies. The current legislative procedure approved for use in the United Kingdom is the V.I.P.E.R. parade which presents video images to the eyewitness on a sequential basis.

A V.I.P.E.R. parade sequentially presents individual video clips of faces to the eyewitness on a computer screen. The parade can be administered anywhere; in a police station, hospital, or residence at a time convenient to the witness. The judgement style fostered by sequential presentation is not a relative judgement but rather an *absolute* judgement style (Lindsay & Wells, 1985; Memon & Gabbert, 2003; McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Tredoux & Chiroro, 2006; Lindsay & Pozzulo, 1999; Kneller, Memon & Stevenage, 2001; Wells *et al.,* 1998; Clark & Davey, 2005). An absolute judgement style removes the witness's ability to compare across parade members and instead forces the witness to compare each parade member with their memory of the perpetrator before making a decision (Wells, 1984; Lindsay & Wells, 1985; Memon & Gabbert, 2003; Steblay *et al.,* 2001; Wells *et al.,* 1998). As with traditional identity parades, it is the judgement style employed which appears to drive the success of V.I.P.E.R. parades as they are claimed to preserve the hit rate for target present parades associated with simultaneous parades (Lindsay & Wells, 1985; Steblay *et al.,* 2001; Wells *et al.,* 1998; Wells, 2008; Valentine, 2006). Although there is suggestive evidence that V.I.P.E.R. parades slightly reduce correct identifications (Wells, 2008; Memon & Gabbert, 2003; Steblay *et al.,* 2001), there is strong evidence that sequential presentation of identity parades significantly reduces the number of false identifications from target absent parades (Wells, 1984; Lindsay & Wells, 1985; Steblay *et al.,* 2001; Wells *et al.,* 1998; McQuiston-Surrett *et al.,* 2006; Gronlund, 2004).

Not only does V.I.P.E.R. have these psychological benefits for eyewitness performance it has also improved the procedural and economic aspects for operating police forces. These advantages are namely a decrease in resources; admittedly there is a large, but necessary, expenditure to buy the recording suite hardware and ensure appropriate staff training yet after this initial outlay the costs involved are minimal (Slater, 1995; Steblay *et al.,* 2001; Valentine, 2006). With V.I.P.E.R. there are no longer any constraints upon recruiting appropriate foils; there is a large central database of images (currently using approximately 12,000 video clips) which will select a range of foils deemed to match a written description of the suspect, and provide them to the organising officer to select the most suitable for inclusion (The Police and Criminal Evidence Act 1984, Code D, 2008; Valentine & Heaton,

1999; Steblay *et al.,* 2001; Slater, 1995). One of the most important effects of using foils from the database is that the chances of an innocent foil being prosecuted from a false identification are drastically reduced (Valentine & Heaton, 1999) as these foils are known to be innocent. Thus, V.I.P.E.R. parades are seen as fairer than simultaneous parades (Valentine & Heaton, 1999; Valentine, 2006; Steblay *et al.,* 2001). This central database of approximately 12,000 images from across the United Kingdom has also reduced police expenditure on foils attendance. There are widely publicised recording schemes and those individuals who attend are paid a one-off fee of £15. Their image is then added to the database and may be repeatedly used for parade construction across the United Kingdom.

Finally, perhaps the most important added psychological benefit to witness performance is that the retention interval associated with V.I.P.E.R. parades is significantly shorter than that seen with traditional parades, averaging around ten days between incident and recall (Valentine, 2006; Steblay *et al.,* 2001; Slater, 1995; Valentine *et al.,* 2003). While such a delay is still liable to memory decay it should not be as significantly impaired as it would be after ten weeks (Deffenbacher *et al.,* 2004; Ebbinghaus, 1885; Wells & Olson, 2003; Ebbesen & Rienick, 1998; Kemp *et al.,* 2001; Bromby, 2002; Dunning & Stern, 1992; Wells, 1978; Christianson, 1992).

So what is the evidentiary support for each of these two types of identity parade? The literature supporting simultaneous superiority for target present parades is certainly voluminous (Levi, 1998; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Valentine *et al.,* 2007; Steblay *et al.,* 2001; Pike *et al.,* 2000; Valentine, 2006; Wells *et al.,* 1998; Kemp *et al.,* 2001; Memon & Gabbert, 2003; Steblay, Dysart & Wells, 2011) and pervading; from Lindsay & Wells initial 1985 study comparing simultaneous and sequential identity parades there has been a general acceptance that, when the suspect is present in the parade, it is beneficial to employ a simultaneous procedure. This advantage certainly appears to be consistent; Steblay, Dysart, Fulero & Lindsay's 2001 meta-analysis investigated over four thousand data points from nine published and thirteen unpublished studies and reported the simultaneous parade format to elicit a correct identification rate of 50% for target present parades

compared with a rate of just 35% for sequential parades. Although neither of these hit rates is particularly good the data shows that simultaneous parades are 30% better than the sequential parades under these conditions. However, this pattern is reversed when the suspect is not present in the identity parade. Steblay *et al.* (2001) found that an innocent foil is identified, on average, in 27% of simultaneous identity parade compared with a rate of just 9% when a sequential format is employed.

There have been challenges to Steblay *et al.'s* (2001) findings in recent years; Carlson, Gronlund & Clark (2008) designed an experiment to exactly replicate the original research which reported the sequential superiority advantage (Lindsay & Wells, 2001) with only one change to how the foils were selected. The authors assumed that the original research had used a biased line-up which appeared to lead to the designated innocent foil being selected almost as frequently as the culprit in target present parades (43% versus 58%). Carlson *et al.* ensured that the selected foils were a close match in similarity to both their culprit and designated innocent foil so as to present a fair, unbiased parade. Despite hypothesising that their results would yield evidence for the sequential parade advantage, the data revealed there to be no significant difference in the rate of correct rejections for target absent parades. The authors suggested that this is evidence that the sequential superiority effect only becomes evident when the identity parade is biased or when the suspect is presented late in the parade sequence.

In their 2011 meta-analysis, Steblay, Dysart & Wells examined data from 72 studies from 23 labs worldwide totalling 13,000 participants and again reported that the sequential identity parade significantly reduced the number of mistaken identifications from target absent parades whilst the simultaneous format is associated with significantly more correct identifications for target absent parades. However, the authors did re-clarify what they meant by 'superiority'; stating that it is a higher diagnositicity ratio which indicates how much more likely one event is in relation to another. As a result of their meta-analysis, they claimed that the sequential parades had a diagnositicity of 7.72 compared with 5.78 for the simultaneous parades. Thus the evidence shows a clear benefit to the judicial system when

sequential parades are employed for target absent identity parades, although there is a decrease in the number of correct identifications associated with this format when the perpetrator is present in the parade.

This effect has been replicated in more recent studies; Memon & Gabbert's 2003 study again found that sequential parades were associated with fewer identifications, of either suspect or foil, overall. Despite this evident disadvantage, Wells has claimed that Steblay *et al.'s* 2001 meta-analysis suggests that an identification obtained from a sequential parade is twice as likely to be strong evidence of suspect guilt compared to an identification made during a simultaneous identity parade (Wells, 2001) and it is because of this point that the sequential identity parade is recommended by over 80% of the published authors (Kassin *et al.,* 2001; McQuiston-Surrett *et al.,* 2006).

While there is a growing body of evidence to support this notion of Sequential Superiority (Lindsay & Wells, 1985; Levi & Lindsay, 2001; Levi, 1998; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Wells, 2001; Wells *et al.,* 1998, Carlson *et al*., 2008) it is also important to note that there are a range of limitations associated with this theory and some issues regarding the validity of the extant literature (Steblay *et al.,* 2001; Wells, 1993; 2001; McQuiston-Surrett *et al.,* 2006). It is recognised by the initial proponents of the sequential procedure that the current literature has not definitively defined the exact conditions under which sequential identity parades can be said to be resolutely superior over simultaneous parades (Wells, 1993; 2001; 2008; Levi & Lindsay, 2001), an issue still under contention in the most recent papers (Memon & Gabbert, 2003; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Valentine *et al.,* 2007; Valentine, 2006; Wells *et al.,* 1998). The primary debate appears to stem from the, as yet unspecified, origins of the sequential superiority effect: is it driven by the specific type of judgement strategy employed or is it a shift in recognition criteria? (McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Gronlund, 2004; Meissner *et al.,* 2005; Kneller *et al.,* 2001).

The prevalent hypothesis has been that of a differing judgement style fostered by each identity parade format; relative judgements with simultaneous parades and absolute judgements with sequential parades (Wells & Lindsay, 1984; Wells *et al.,* 1998; Steblay *et al.,* 2001; Memon & Gabbert, 2003; McQuiston-Surrett *et al.,* 2006). Such terminology was first used by Wells (1984) yet its evidential basis is derived purely from unreliable self-reports from participants and observational analysis that, in target absent simultaneous parades, witnesses typically lean towards selecting another foil rather than electing to reject the entire parade (Wells, 1993). Of the 54% of witnesses who correctly identified the suspect in the target present parade 43% went on to make a false identification in the target absent parade; a step frequently termed *target-to-foil shift* (Clark & Davey, 2005) and is claimed to be indicative of relative judgements whereby, in the absence of the suspect, the next best matching foil is identified (Clark & Davey, 2005; Steblay *et al.,* 2001, Lindsay & Wells, 1985, McQuiston-Surrett *et al.,* 2006, Gronlund, 2004). This is particularly an issue for V.I.P.E.R. identity parades where the 'closeness of match' between the suspect and the foils is much close than in live identity parades. In their 1999 study, Valentine & Heaton assessed the fairness of V.I.P.E.R. in comparison to the traditional, live format. If an identity parade is completely fair then each individual within the parades has an equally chance of being selected (Valentine, 2006). Working from a parade size of 1 suspect and 8 foils, this equated to the suspect being selected by 11% of participants. Valentine and Heaton tested both the V.I.P.E.R. format and the simultaneous format for fairness using 216 participants. Their results illustrated that the suspect was identified by 25% of participants, indicating that they were selected more frequently than the 8 foils; thus suggesting that the simultaneous parades were not fair and were biased towards the suspect. By comparison, the suspect was selected by 15% of participants who viewed the V.I.P.E.R. identity parade. Further analysis revealed that there was a significant difference in fairness between the V.I.P.E.R. and simultaneous parade types; that V.I.P.E.R. was significantly fairer to the suspect than the simultaneous format. Most importantly however, although the V.I.P.E.R. parade format was associated with a greater probability of the suspect being selected by the mock witnesses, this was not significantly different from the expected chance levels (15% versus 11%). The results of this study provide firm evidence that the V.I.P.E.R. identity parade is

significantly fairer to suspects and significantly less biased compared to the simultaneous format.

As sequential identity parades are typically associated with higher accuracy rates for target absent conditions and this format is claimed to rely upon an absolute judgement style this appears to be further support for the judgement style hypothesis. It is this evidence and that from other corroborating studies (Kneller *et al.,* 2001; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Gronlund, 2004) which have led to the judgement style hypothesis becoming widely accepted within the literature.

Yet there are some indications that the target-to-foil shift does also occur for sequential parades when the suspect is absent from the identity parade. In their 2005 study, Clark & Davey investigated this phenomenon in both target present and target removed parades. They hypothesised that the effect would occur for simultaneous parade but that it would be diminished or absent from the sequential parade because of the diminishing use of relative judgement styles with this format. While their data further corroborated the overall advantage for sequential identity parades they also found evidence that the target-to-foil shift was equal for both simultaneous and sequential parades. Such a result suggests that when the target is removed from the suspect witnesses are equally as likely to gravitate towards identifying an innocent foil regardless of the parade format employed. This data is of vital importance to understanding the underlying cognitive processes of eyewitness judgement styles and is a clear indicator that further research is necessary to explore exactly how eyewitness identification is elicited cognitively and to determine the exact judgement style employed by witnesses in both identity parade formats.

There have also been some claims that the sequential superiority effect may be underpinned by a shift in recognition criteria rather than the judgement style used by witnesses (McQuiston-Surrett *et al.,* 2006; Meissner *et al.,* 2005; Ebbesen & Flowe, 2002). The basis of this hypothesis arises from the procedural differences between the simultaneous and sequential format. In the format currently recommended by the research community,

sequential parades are essentially a forced choice task: the images are presented sequentially but, crucially, the eye-witness must make a decision before the parade can proceed (Levi & Lindsay, 2001; Steblay *et al.,* 2001; Wells, 2001; Lindsay & Wells, 1985; Lindsay & Pozzulo, 1999; Valentine *et al.,* 2007). Thus, the eyewitness is required to either accept the current image as the suspect or reject it entirely and view the next image. It is proposed then that, as there is no possibility of comparing parade members, witnesses must evaluate whether the current image is sufficiently similar enough to their memory of the suspect or whether they should reject this image in case a subsequent image proves to be a stronger match (Valentine *et al.,* 2007; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Gronlund, 2004; Memon & Gabbert, 2003; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999). Clark & Garvey (2005) found some evidence of there being a shift in recognition criteria: when the suspect was removed from the identity parade some witnesses downgraded their recognition criteria in order to make an identification. However, while this is valuable evidence that recognition criteria *can* be altered between parade formats it is suggesting a converse effect to that observed in other studies. Many of these studies have suggested that the sequential format leads eyewitnesses to become more conservative in their identifications (Ebbesen & Flowe, 2002; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Kneller *et al.,* 2001; Memon & Gabbert, 2003) which would account for the widespread effect of reducing both false positive identifications *and* correct identifications.

Some corroborating data for this can be found in Memon & Gabbert's 2003 study where they report that their participants were more likely to not make an identification during sequential target present parades. Ebbesen & Flowe (2002) also concluded that the use of sequential identity parades encourages witnesses to adopt a more conservative recognition criterion and, in their 2005 study, Meissner, Tredoux, Parker & MacLin employed a series of experiments based on signal detection theory, results of which appear to confirm that sequential identity parades are associated with a more conservative criterion.

This judgement style vs. criterion shift debate remains on-going and further research is certainly needed to resolve this issue but there is a more pressing concern regarding the

procedural validity of the majority of simultaneous vs. sequential studies. As mentioned previously, the sequential format employed across many of the current studies adheres to a truly sequential procedure as advocated by many researchers (McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Lindsay & Wells, 1985; Levi, 1998; Lindsay & Pozzulo, 1999; Levi & Lindsay 2001; Kassin *et al.,* 2001; Wells *et al.,* 1998; Valentine *et al.,* 2007; Valentine, 2006; Wells & Olson, 2003). This strict procedure requires that the images be sequentially presented but that the eyewitness *must* make a decision; accept or reject. If the current image is identified as the suspect then the parade is immediately halted and no further images are shown. If the witness rejects the current image the subsequent images are shown. The witness must not be told how many images they will see and once an image has been rejected it cannot be accepted at a later time (Steblay *et al.,* 2001; Lindsay & Pozzulo, 1999; Wells *et al.,* 1998; Memon & Gabbert, 2003; McQuiston-Surrett *et al.,* 2006; Valentine, 2006). Finally, double blind administration is also required under this strict format to escape any possible bias (Lindsay & Pozzulo, 1999; Wells, 1993; 2001; Wells *et al.,* 1998; Levi & Lindsay, 2001; Valentine *et al.,* 2007; Valentine, 2006; Valentine & Heaton, 1999; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Kassin *et al.,* 2001); the provisions of PACE Code D allow for double-blind presentation.

This strict sequential procedure represents a problem, certainly to UK researchers, as the V.I.P.E.R. identity parade procedure is bound by the limitations of Code D of The Police & Criminal Evidence Act (1984), which explicitly states that all images contained within a V.I.P.E.R. parade be shown to the witness at least twice *before* they can make an identification or reject the parade (Valentine, 2006; The Police and Criminal Evidence Act 1984, Code D 2008). Additionally, the Act permits individual images to be examined in more detail by the witness if so required (The Police and Criminal Evidence Act 1984 Code D 2008). With regards to double blind administration this is required and executed, as many British police forces have a team of dedicated V.I.P.E.R. officers who each administer different steps of the parade to ensure double-blind presentation.

This procedural difference between the 'sequential parades' of the literature and the UK V.I.P.E.R. is a key point in this thesis; much of the existing literature laments that the precise conditions under which sequential superiority is achieved remains unknown despite extensive research. This is primarily because of the range of factors that have been investigated and the divergent methodologies employed; there has been little effort to standardise exactly what is meant by a 'sequential' identity parade. The large majority of the extant literature has focussed on the benefits of a strict sequential procedure and so the results of such research cannot be fully generalised to the UK V.I.P.E.R. parade; for V.I.P.E.R. is not a strict sequential procedure. We have already noted that UK legislation stipulates that witnesses must view the entire parade twice before being permitted to make and identification. Furthermore, V.I.P.E.R. parades make allowances for witnesses to re-visit any of the parade individuals repeatedly, to pause and examine the image in detail and to allow witnesses to change their minds about a previous identification (The Police & Criminal Evidence Act (1984), Code D (2008)). So it appears that the sequential, video-based identity parade format of V.I.P.E.R. has yet to be investigated in its current British format.

Thus, the theories based upon the extant literature, while useful to a degree, are based upon a procedure which would not be permissible under UK law. V.I.P.E.R. is clearly an entirely different procedure to that employed by many research papers which have tested the sequential advantage. To date, there has been only one published study which has tested V.I.P.E.R. as it is currently applied in the United Kingdom. This paper investigated whether strict pre-parade instructions or moving images increased the reliability of the sequential V.I.P.E.R format and is reviewed in more detail later in this thesis.

It is therefore evident that there is an urgent need for investigations of the effects provoked by the V.I.P.E.R. format to determine whether a similar effect to that widely reported in the literature is elicited: that simultaneous parades are superior for target present identity parades while sequential is superior for target absent parades.

Special Measures for Vulnerable Witnesses

The next stage for many eye-witnesses is the progression of their case to trial; this experience is recognised by many researchers and witness support groups to be very traumatic, especially if they have been subjected to intimidation (Tarling, Dowds & Budd, 2000; Fyfe & McKay, 2000; Editorial Criminal Law Review Editorial, 2004). In an effort to improve the witness experience of this stage of the criminal justice system, several consultations were begun to investigate the causes of witness anxiety and how this could be ameliorated to the benefit of both witnesses and the justice system. It is patently clear that if a witness is exceptionally nervous about testifying in court the quality of their testimony will be severely compromised (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003; Golding *et al.,* 2003; Goodman *et al.,* 1998; Swim *et al.,* 1993) which in turn may affect the jurors' perceptions of that witness and could significantly influence the outcome of the trial (Goodman *et al.,* 1998; 2004; Orcutt *et al.,* 2001; Richards, Morris & Richards, 2008; Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003).

Research investigating the effects of different factors upon jurors' perceptions and decision-making processes face a difficult task. By their very nature, juries and their deliberations are highly confidential; this means that researchers cannot utilise actual deliberating jury groups for their studies. Certainly after the conclusion of the trial the jurors can be approached and asked to recount their memories of the events and processes that occurred during deliberation but these types of self-report measures are often biased and do not yield a complete picture of the true nature of deliberations (Tinsley, 2001; Zander & Henderson, 1993; Jackson, 1996). Other researchers have attempted to use 'shadow juries' as their participants (McCabe & Purves, 1974). These 'shadow juries' involve recruiting a group of participants and exposing them to exactly the same trial evidence as the actual jury. To this end, the shadow jury group are required to attend court for each and every day of the selected trial to sit in the public gallery and listen to the testimony and view the evidence, before being committed to deliberations as the actual trial jury would be. Of course, the issues involved in such a research design are apparent; ensuring that every participant

attends court every day, requiring them to commit to a study which lasts several hours for an undefined period and then requiring them to deliberate on what they have hear during that time. It is not surprising that this type of jury research is rare and in response to these issues, jury researchers have been compelled to devise an alternative methodology for studying jurors' perceptions and decision-making. This has become known as the 'mock juror paradigm'.

The mock juror paradigm involves recruiting participants eligible for jury duty and exposing them to a mock, or simulated, trial. This methodology of jury research has been widely adopted (Goodman *et al.*, 1998; 2004; Orcutt *et al.*, 2001; Swim *et al.*, 1993, Davies & Noon, 1991; Nisbett & Ross, 1980; Landstrom *et al.*, 2005, 2008; Hope, Greene, Memon, Gavisk & Houston, 2008; Tinsley, 2001; Pennington & Hastie, 1990; Penrod & Cutler, 1995; Quas, Thompson & Clarke-Stewart, 2005; Pozzulo, Crescini & Panton, 2008; Bornstein, 1999; Nemeth, 1981; Kramer, Kerr & Carroll, 1990; Hastie, Penrod & Pennington, 1983; Carlson & Russo, 2001; Blamey, McCarthy & Smith, 2000). Typically, the mock juror participants are recruited from the student body population, which has prompted concern from some researchers (Bray & Kerr, 1982; Casper, Benedict & Perry, 1989; Goodman, Golding, Helgeson, Haith & Michelli, 1987) claiming such data cannot be generalised to the wider community, but several large-scale reviews and meta-analyses have revealed that data derived from such samples does not have limited generalisability (Bornstein, 1999; Pozzulo *et al.*, 2008). Much of the research employing the mock juror paradigm does differ from actual trials on one major point; the mock jurors frequently work in isolation (Goodman *et al.*, 1987; Landstrom *et al.*, 2005; 2008; Goodman *et al.*, 1998; 2004; Orcutt *et al.*, 2001; Swim *et al.*, 1993; Nisbett & Ross, 1980). Many researchers opt to run participants on an individual basis, as it is cheaper and less time consuming. Few studies have elected to run a full-scale mock trial, simply because of the demand it places on resources. It is a difficult task to arrange a large group of participants to attend at a specific time. Last minute drop-outs and no-shows can lead to the whole group being cancelled; juries must adhere to a standard size, if numbers are diminished then the mock trial cannot proceed. Despite the differences, reviews of the findings from over 45 years of mock jury research has advocated

these simulation studies as providing researchers with better opportunities for experimental control and increased freedom to test theories and procedural variations which would not be permitted in the courtroom (Pozzulo *et al.*, 2008; Bornstein, 1999; Diamond, 1997). Most importantly, the conclusions drawn from these mock juror studies may be readily generalised to the wider populations (Bornstein, 1999; Diamond, 1997).

Jurors are notoriously susceptible to a host of influencing factors ranging from order effects (Bruine de Bruin, 2005; Bruine de Bruin & Keren, 2003; Kassin & Wrightsman, 1979; Walker, Thibault & Andreoli, 1979; Pennington & Hastie, 1992; Carlson & Russo, 2001), attractiveness of witness and defendants (Sigall & Ostrove, 1975; Castellow, Wuensch & Moore, 1990; Darby & Jeffers, 1988; Ahola, Christianson & Hellstrom, 2009), jury instructions (Hastie, Penrod & Pennington, 1983; Daftary-Kapur, Dumas & Penrod, 2010), speech styles and body language (Wheatcroft, Wagstaff & Kebbell, 2004; Aron & Rosner, 1998; Boccaccini, 2002; Boccaccini, Gordon & Brodsky, 2005; O'Barr, 1974; Conley & O'Barr, 1998; Gibbons, 1995; Conley, O'Barr, & Lind, 1978; Erickson, Lind, Johnson, & O'Barr, 1978; Lind, Erickson, Conley, & O'Barr, 1978; O'Barr, 1982; O'Barr & Conley, 1976) and group dynamics (Hastie *et al.,* 1983; McCoy, Nunez & Dammeyer, 1999; Stasser, Kerr & Bray, 1982; Blamey, McCarthy & Smith, 2000; Devine, Clayton, Dunford, Seying & Pryce, 2001). Manipulation of these factors has been shown to alter jurors' perceptions and understanding and to change the outcome of a trial. Witness demeanour is another factor which can exert a significant influence upon jurors' perceptions of the witness and the reliability of their testimony yet this remains an under-researched area. One such study by Golding, Fryman, Marsil & Yozwiak (2003) investigated the effect of emotion and demeanour upon mock jurors' perceptions of child witnesses. When the child witness remained calm and emotionless during testimony they were perceived as significantly less credible and reliable than the child witness who displayed some emotion (crying). Yet, too much emotion, or hysteria, is equally likely to reduce perceived witness credibility. There is some additional support for this effect (Regan & Baker, 1998) and while the effects of witness demeanour, particularly involving an adult witness, upon juror perceptions remains uncertain, the effect of high emotions and anxiety can be readily extended from these earlier studies.

If a witness is so severely anxious about testifying it is expected that this will have a significant, detrimental effect upon the quality of their testimony (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003; Golding *et al.,* 2003; Richards, Morris & Richards, 2008). A highly nervous witness is more likely to answer questions in a disjoined manner, with frequent outbursts of emotion, poor eye contact and 'fidgety' body language (Aron & Rosner, 1998; Boccaccini, 2002; Conley & O'Barr, 1998; Gibbons, 1995; Conley, O'Barr, & Lind, 1978; Erickson, Lind, Johnson, & O'Barr, 1978; Lind, Erickson, Conley, & O'Barr, 1978; O'Barr, 1982; O'Barr & Conley, 1976). This is significant; relaxed, assured body language and maintaining strong, but not intimidating, eye contact have been proven to be heavily relied upon by jurors as a reliable indicator of witness credibility and possibly deceit (Penrod & Cutler, 1995; Boccaccini, 2002; Boccaccini *et al.*, 2005; Tetterton & Warren, 2005; Aron & Rosner, 1998). Poor performance in these areas may lead jurors' to discount valuable testimony purely because of a witness's nerves (Golding *et al.,* 2003; Regan & Baker, 1998; Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003) Therefore, after an extended period of consultation with judicial professionals, witness support services and the general public a range of 'Special Measures' for vulnerable witnesses was outlined and later incepted into law in Scotland under The Vulnerable Witnesses (Scotland) Act 2004. These 'Special Measures' were designed to improve the witness experience when at court by circumventing the legislative right of defendants to face their accuser thus removing the primary cause of witness anxiety yet preserving the basic rights of the accused (Birch, 2000; Tausz & Ellison, 2005; Hoyano, 2001; The Vulnerable Witnesses (Scotland) Act 2004).

The Special Measures are not a new initiative; many of the testimony methods included under the Act have been routinely available to child witnesses, especially those linked to sexual abuse cases (Hamlyn *et al.,* 2004; Richards, Morris & Richards, 2008; Goodman *et al.,* 1998; 2004; Swim *et al.,* 1993, Landstrom, Granhag & Hartwig, 2005; 2008). The major change in legislation arises from the extension of the permissible criteria for making an application to use the Special Measures. Since 2004 the full range of Special Measures has been available for adult witnesses; providing they meet specific criteria (The Vulnerable

Witnesses (Scotland) Act, 2004, Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003); a documented history of intimidation in relation to their testimony and/or severe stress and anxiety to the detriment of their evidence. Additionally, the adult witness applying for permission to use the Special Measures is required to have attended a V.I.P.E.R. identity parade and made a positive identification (Richards, Morris & Richards, 2008; The Vulnerable Witnesses (Scotland) Act, 2004). Thus the provisions of the Act ensures that the Special Measures are not available to all adult witnesses; but only upon application to the trial judge who decides whether the witness's testimony is likely to be significantly impaired by appearing in court without the use of the requested measure.

The requested Special Measure may be any of a range allowed under The Vulnerable Witnesses (Scotland) Act including in-court support from The Witness Service, submission of previous statements as evidence-in-chief and the use of pre-recorded video evidence, in addition to more familiar methods such as use of a screen and live CCTV links.

However, while the public consultations were extensive there was no attempt to empirically investigate the effect of the Special Measures upon jurors' perception of vulnerable adult witnesses and if this alteration then influences the trial verdict. The existing literature has gone some way towards answering this question however; many of these studies have used child witnesses as the protagonist (Swim *et al.,* 1993; Davies & Noon, 1991; Goodman *et al.,* 1998; 2004; Orcutt *et al.,* 2001; Landstrom *et al.,* 2005; 2008). This is not entirely unexpected; the inception of the Special Measures for use by adult witnesses is relatively recent while child witnesses have had access for a number of years. While this lack of adult data may be due to a publishing delay we may perhaps draw some tentative inferences from studies which have employed child witnesses.

Most of the relevant literature has focussed on one of the more frequently used Special Measures, live CCTV link, and the effects it exerts upon mock jurors. One of the first, and major, studies investigating the effect of CCTV link was Goodman, Tobey, Batterman-Faunce, Orcutt, Thomas, Shapiro & Sachsenmaier (1998) revealed a small yet significant

effect upon mock jurors' perception when the witness testified via a CCTV link. 186 children, 'witnesses' and over 1200 mock jurors were recruited for the study. The child was allowed to build a rapport with a 'baby-sitter' over 2 sessions which culminated in the child being asked to place stickers on either exposed body parts (toes, upper arm) or covered body parts (shirtsleeve, socks) thus forming a 'defendant guilty' and 'not guilty' condition. The third session was a mock trial where half of the children testified without a CCTV link whilst the others used this Special Measure; the children were questioned about the earlier sessions with the baby-sitter while the mock jurors observed the trial. Jurors were asked to rate the witness's believability, accuracy, intelligence and suggestibility and then to deliver a verdict upon the case. What Goodman *et al* found was using a CCTV link to testify was associated with a small, yet significant, negative bias in jurors' perceptions of the witness; the children who testified in this manner were rated as less believable, less intelligent and more suggestible than the children who testified in an open court with no Special Measures. However, jurors did rate them as more accurate than the witnesses who testified in an open court. Furthermore, the negative bias for believability was not linked with a significant effect upon defendant conviction rates; the jurors were equally as likely to convict regardless of the testimony presentation method although the ratings of defendant guilt were marginally greater under the CCTV condition but this was not significant.

In an earlier study Swim, Borgida & McCoy (1993) focussed on a different aspect of the issue of in-court Special Measures; whether the defendant's rights are compromised while also assessing jurors' reactions to child witnesses who chose to testify via a pre-recorded deposition. Their results indicate a similar pattern: jurors who viewed the pre-recorded evidence condition perceived the defendant less positively than the jurors who viewed the witness testify in an open court; although this was not significant and did not exert an associated increase in conviction rates. Equally, the jurors reported no significant emotional responses to observing the child witness testify via pre-recorded evidence and did not view them any more sympathetically.

In more recent research, Landstrom, Granhag & Hartwig (2005; 2008) of the University of Gothenburg have expanded these findings. In their 2005 study, they investigated the effects of witnesses appearing live versus on video upon mock jurors' ratings of witness veracity, their memory and their perceptions. 12 participants were recruited to serve as 'witnesses' and were exposed to a staged event. The event was a car accident in a parking lot, performed by two professional stuntmen; one drove a car through the parking lot whilst talking on a mobile phone. Whilst attempting to turn left, the driver moved too far right and hit the other stuntman who was on a bicycle. The cyclist lay still on the ground 'injured' while the driver exited the car to check on him; the cyclist was 'injured'. The 'witnesses' had been instructed to wait at a specific location in parking lot and the entire event lasted less than five minutes. The witnesses were then split into two groups; one told the truth about the incident whilst the others were incited to lie during testimony for a financial benefit (600 USD). All witnesses were given 5 minutes to prepare for their interview with the police. Three weeks later, all witnesses returned for a court hearing. Each witness was required to testify which took the form of free recall of the event followed by 11 specific questions. All witnesses were viewed during their testimony by 4-7 mock jurors; this was the 'live' testimony condition. The testimony was also recorded for the 'video' testimony condition and was shown to an equal number of mock jurors as viewed the 'live' testimony. The results of Landstrom *et al.'s* experiments suggested that the witnesses viewed live were rated as significantly more eloquent and more pleasant and more honest. Furthermore, the live witnesses were viewed with greater leniency that those viewed via videotape. The authors suggested that these findings were linked to the physical proximity to the witness; that the live witness was viewed more sympathetically than the witness viewed via videotape and was also less likely to be critically evaluated. The only limitations of the Landstrom *et al* study are that they did not ask their mock jurors to provide a verdict upon the case and so we cannot gauge what effects this change in jurors' perceptions exerts upon the likelihood of a guilty verdict, and that the adult witnesses were not 'vulnerable'.

Sara Landstrom's research in this area of CCTV and videotaped evidence as alternative testimony presentation methods are the result of the continuing debate in Sweden over the

adoption of Special Measures into Swedish Law. Special Measures are currently not permitted in Sweden and the implementation of Special Measures in the United Kingdom, and in Scotland in particular, is being closely monitored to assist Swedish researchers and legal practitioners in their push for such a reform to their own legislation.

The Goodman *et al.* (1998) and the Swim *et al.* (1993) studies along with the Landstrom *et al.* (2005) paper anticipated that the use of a CCTV link to testify would be associated with mock jurors' adopting a more pro-prosecution stance; that they would perceive the child witness as more credible and believable than a child witness who testified in an open court and would therefore be more likely to convict the defendant. This reasoning stems from the explanation that jurors reach when viewing a witness testifying via Special Measures: that there must be a *reason* why the witness is too nervous to come into the courtroom thus the defendant must have done *something* to warrant this emotional response (Regan & Baker, 1998; Golding *et al.,* 2003).

So why does this effect not occur in the extant literature? One theory proposed by Davies & Noon (1991) is that Special Measures which remove the witness from the direct observation of the jurors are associated with reduced immediacy and exerts a weaker emotional impact. We have previously mentioned data which corroborates Davies & Noon's theory; Swim *et al* (1993) found that their mock jurors did not respond any more empathetically to the CCTV link witness while the participants of Goodman *et al.'s* (1998) study perceived the testimony of CCTV link witness as less reliable. This theory is similar to an effect first proposed in 1980 by Nisbett & Ross; the *vividness effect*. Their participants viewed both live and videotaped testimonies and were intensively questioned about their perceptions after each. Nisbett & Ross reported that live testimonies were associated with increased vividness which led mock jurors to pay more attention to them, to consider live witnesses more credible and more memorable than the videotaped testimonies. Nisbett and Ross clearly defined why they believed this effect to hold; they claimed that vivid testimonies are characterised by increased emotional impact or intensity, increased use of imagery and, most relevant to the use of Special Measures, by proximity of the witness to the jurors. The

more proximate the witness is to the juror, the more vivid the testimony, and the more credible the witness appears.

One further area we can perhaps draw inferences to explain this effect is from the extensive research into witness body language and, to a lesser extent, speech styles and how this may be being interpreted by jurors (Boccaccini, 2002; Aron & Rosner, 1998; O'Barr, 1974; Penrod & Cutler, 1995; Tetterton & Warren, 2005; Landry & Brigham, 1992; Orcutt *et al.,* 2001; Landstrom, Granhag & Hartwig, 2005; 2008; Conley & O'Barr, 1998; Gibbons, 1995; Conley, O'Barr, & Lind, 1978; Erickson, Lind, Johnson, & O'Barr, 1978; Lind, Erickson, Conley, & O'Barr, 1978; O'Barr, 1982; O'Barr & Conley, 1976). Much of this research originates from the United States where the preparation of witnesses for trial is a large component of their judicial system; it is common for both prosecution and defence to spend a significant amount of time, and money, coaching witnesses on the most effective manner of delivering their testimony (Boccaccini, 2002; Aron & Rosner, 1998; Boccaccini *et al.,* 2005).

It is widely recognised that jurors base their determinations of witness credibility and honesty on observations of their behaviour during testimony (Boccaccini, 2002; Aron & Rosner, 1998; Penrod & Cutler, 1995; Tetterton & Warren, 2005; Landry & Brigham, 1992; Boccaccini *et al.*, 2005). US witnesses are routinely coached in effective testimony delivery (Aron & Rosner, 1998; Boccaccini, 2002) including confident body movements: using illustrator gestures to reinforce their point; avoid fidgeting on the stand and extremes of emotion; appear relaxed and confident and maintain assertive, but not aggressive, eye contact with the jury. Clearly when a witness testifies in an open court, with no Special Measures, the juror can easily observe the behaviour of the witness directly. Similarly, when a screen is employed as a Special Measure this does not interfere with jurors' direct observation of the witness as it is designed purely to shield the witness from the defendant's view (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003; The Vulnerable Witnesses (Scotland) Act, 2004). Yet when we progress to the live CCTV link and Pre-recorded Video Evidence Special Measures the negative effect upon jurors' perceptions is

perhaps reasonable. If we first assess the live CCTV link method it is immediately obvious that the jurors' observation becomes *indirect*; as the witness does not appear before the jury but via a television link from a separate location (The Vulnerable Witnesses (Scotland) Act, 2004). In an open court the jury can see much more of the witness, and therefore more of their behaviour, whereas the witness can be seen only from the shoulders up when testifying via a CCTV link. In addition, during the link the witness cannot see the jury directly and thus cannot deliver effective eye contact, so important to jurors' perceptions and detection of deceit (Landry & Brigham, 1992; Boccaccini, 2002; Aron & Rosner, 1998; Mehrabian, 1981). Despite this obstacle the CCTV link Special Measure does permit some indirect observation of the witness's body language when under pressure, particularly under examination by a challenging defence lawyer (Wheatcroft *et al.,* 2003; Boccaccini, 2002; Aron & Rosner, 1998).

When a vulnerable witness elects to testify via pre-recorded video evidence, or statements, it decreases their proximity from the jury and from the jurors' close scrutiny whilst also eliminating the adversarial nature of a live trial (Wheatcroft *et al.,* 2004; Swim *et al.,* 1993; Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003; Birch, 2000; Regan & Baker, 1998). The use of pre-recorded video evidence requires early action from both prosecution and defence lawyers to attend an interview session with the witness. Frequently the video statement may be recorded many months in advance of the actual trial thus the witness may not be required to attend the trial and confront the accused (The Vulnerable Witnesses (Scotland) Act, 2004). The use of pre-recorded video evidence is associated with a significant decline in witness anxiety (Criminal Law Review Editorial, 2004; Richards, Morris & Richards, 2008) which enables such witnesses to deliver their testimony in a relaxed manner without excessively strong emotional responses. Furthermore, the examination and cross-examination is completed in a less adversarial manner, further reducing the witness's emotional response. While these alterations to traditional testimony procedure are certainly beneficial to the vulnerable witness these same differences may be what prompts the jurors to perceive the witness in a negative light (Landstrom *et al.,* 2005;

Goodman *et al.,* 1998; 2004; Orcutt *et al.,* 2001; Davies & Noon, 1991; Nisbett & Ross, 1980).

It is not unreasonable to speculate that the negative reflection upon a witness using pre-recorded video evidence is a derivative extension of Davies & Noon's immediacy and emotional impact theory; if the witness is relaxed and composed throughout their recorded statement this may trigger jurors' deception detection as they perhaps expect a more emotional response from the witness. This is an appropriate conclusion; the use of pre-recorded video evidence is typically reserved for the most vulnerable of witnesses in the most serious offences (The Vulnerable Witnesses (Scotland) Act 2004; Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003; Richards, Morris & Richards, 2008; Swim *et al.,* 1993) such as prolonged sexual abuse, hate crimes and child abuse (Swim *et al.,* 1993; Goodman *et al.,* 1998; 2004; Orcutt *et al.,* 2001; Landstrom *et al.*, 2005; Richards, Morris & Richards, 2008; Elliott, 1998).

Thus jurors would anticipate an emotional response from the witness as they are questioned about a traumatic, recent experience regardless of the presence or absence of the defendant (Golding *et al.,* 2003; Regan & Baker, 1998). If this emotional and/or behavioural response is lacking or weak then jurors' may assume that the witness's testimony has been embellished or is entirely fabricated; particularly in trials where the victim is the *only* prosecution witness (Richards, Morris & Richards, 2008; Hamlyn *et al.,* 2004; Davies & Noon, 1991; Orcutt *et al.,* 2001; Goodman *et al.,* 1998; 2004; Criminal Law Review Editorial, 2004).

This is completely the opposite effect of that hypothesised by Swim *et al.* (1993) where it was postulated that the jurors would tend towards empathy with the vulnerable witness and rationalise that they were using Special Measures to testify for a valid reason, such as extreme fear of the defendant (Goodman *et al.,* 1998; Hoyano, 2001; Birch, 2000). The results from the available literature suggests a direct link between the immediacy and

vividness of the witness' testimony and how favourable they are then perceived by the jurors; as the witness becomes increasingly distal from the courtroom and the jury, they are viewed as progressively less credible and accurate (Goodman *et al.,* 1998; 2004; Orcutt *et al.,* 2001; Swim *et al.,* 1993; Landstrom *et al.,* 2005). This is valuable corroborative evidence for Davies & Noon's (1991) theory of immediacy and emotional impact and Nisbett & Ross' (1980) vividness effect and strongly suggests that permitting vulnerable witnesses to testify via remote Special Measures may be exerting a negative effect upon jurors' perceptions of witness credibility (Landstrom *et al.,* 2005; Goodman *et al.,* 1998; 2004; Orcutt *et al.,* 2001) and therefore may further affect jury deliberations and verdict (Landstrom *et al.,* 2005; Davies & Noon, 1991; Nisbett & Ross, 1980, Landstrom, Granhag & Hartwig, 2007).

It is this issue which has prompted furious claims from defence lawyers and human rights charities that defendants' rights are being infringed under the provisions of The Vulnerable Witnesses (Scotland) Act and its 1999 counterpart in England & Wales; The Youth Justice & Criminal Evidence Act (Criminal Law Review Editorial, 2004; Birch, 2000; Hoyano, 2001; Tausz & Ellison, 2005). The legal argument behind these claims is that giving permission to a vulnerable witness to testify using Special Measures is an implicit admission of defendant guilt *before* the trial has even begun (Goodman *et al.,* 1998; 2004; Swim *et al.,* 1993; Birch, 2000; Tausz & Ellison, 2005; Hoyano, 2001), infringing both the right to presumed innocence and the right to confront their accuser as prescribed under Article 6 of The European Convention on Human Rights. If the jury presumes defendant guilt solely from a witness using a Special Measure to testify this is likely to attract a greater rate of conviction (Tausz & Ellison, 2005; Hoyano, 2001; Birch, 2000; Richards, Morris & Richards, 2008). However, a legal appeal to the House of Lords (Tausz & Ellison, 2005) was rejected and it was stated that Special Measures were not contrary to defendants' rights, a conclusion further supported by Hoyano's (2001) review. Further corroborating data, which may be of more value, is that which has been discussed previously. The child witness studies of Goodman *et al.* (1998; 2004) and Swim *et al.* (1993) and the adult witness study conducted by Landstrom *et* al (2005) all suggest that the use of Special Measures does exert a negative

effect but largely upon jurors' perceptions of the witness although there is also some weak evidence of a negative bias towards defendants (Swim *et al.,* 1993). Both of Goodman *et al.* (1998: 2004) and the Swim *et al.* studies though report that there is no link between the negative bias towards the witness, and the defendant, and likelihood of conviction. Therefore, it is indicative that the use of Special Measures does not infringe defendants rights under Article 6 of the European Convention on Human Rights and certainly does not serve to increase the possibility of conviction; if anything the extant literature suggests that the defendant may benefit from the use of Special Measures as they prompt jurors to view vulnerable witnesses and less credible and reliable (Landstrom *et al.,* 2005; Goodman *et al.,* 1998; 2004; Orcutt *et al.,* 2001; Davies & Noon, 1991; Swim *et al.,* 1993; Nisbett & Ross, 1980).

So how widely used are the Special Measures, especially for *adult* vulnerable witnesses? A large scale review of the Scottish system has recently been published (Richards, Morris & Richards, 2008) and examined the implementation levels across the first three years of the Act's inception. This review was conducted on behalf of the Scottish Executive and has not been subjected to peer review; however, it is the first report of the statistics of Special Measures implementation for adult vulnerable witnesses. Adult witnesses were only permitted to use the Special Measures from April 2006 and so the review covers only the first 8 months of adult use until December 2006. Over 50 applications by adult vulnerable witnesses were made during this 8 month period and screens were the most frequently requested by adult witnesses.

Live CCTV links are more commonly requested in High Court trials and are especially popular with children. Richards, Morris & Richards (2008) review also reports that it appears that adult witnesses are moving away from screen use and towards CCTV links. However, despite a system wide initiative to improve the awareness and availability of Special Measures to vulnerable witnesses, Richards, Morris & Richards (2008) reported that some Special Measures were unavailable in some courts three years after initial implementation although applications for Special Measures were very rarely refused by the presiding judge.

Of the 787 vulnerable and/or intimidated witnesses identified by the review 42.2% used a live CCTV link while 22.5% requested the use of a wooden screen. Pre-recorded video evidence was requested and permitted in only a single case in Scotland during the three year review, although it is suggested that the poor availability of recording suites and image quality may be a contributory factor to the poor level of implementation. 77.9% of witnesses identified as 'vulnerable and/or intimidated' were female and 54.7% of these were also the victim in the case.

Richards, Morris & Richards (2008) also questioned adult vulnerable witnesses after they had completed their testimony about their experience in the judicial system. Vulnerable witnesses widely reported that there was a distinct lack of support and information in the time between investigation and trial. Furthermore, many still feel that witnesses are treated poorly during their time at court with frequent contact with the defendant or their associates. Finally, many witnesses claim that they would not be willing to testify in court again unless they were permitted Special Measures (Criminal Law Review Editorial, 2004; Hamlyn *et al.,* 2004).

Richards, Morris & Richards (2008) review gives a clear, valuable insight into the problems associated with vulnerable witnesses' court experience. Although there is a growing feeling within witness groups that the witness experience is improving, there remain doubts about several contentious issues (Tausz & Ellison, 2005; Hoyano, 2001; Landstrom *et al.,* 2005). The data from Swim *et al* (1993) and Goodman *et al* (1998) is indispensable in a currently under-researched area yet it remains an issue that the conclusions from child studies are being widely extrapolated to apply to adult witnesses. The Landstrom *et al.* (2005) is unique it is use of recorded video evidence (although this is of a different format to that used in the United Kingdom) and that it relied upon adult witnesses (who were not 'vulnerable). A child witness elicits significantly more scepticism from jurors as there are valid doubts as to their veracity, memory capabilities, suggestibility and comprehension (Quas *et al.,* 2005; Goodman *et al.,* 1998; Wheatcroft *et al.,* 2004; Regan & Baker, 1998; Golding *et al.,* 2003; Laimon & Poole, 2008). It is therefore of vital importance, as the implementation of Special

Measures for adult vulnerable and/or intimidated witnesses continues to grow, that the effects of such a step upon jurors, witnesses and defendants are closely examined.

Thesis Overview

In recent years there have been several innovative steps taken to improve the quality of eyewitness identification and their later experiences at court. However, there is an apparent short-coming in both the empirical evidence to support the introduction of these reforms and a limit to the extent which the extant literature can be generalised to account for such reforms. This thesis aims to investigate this new route through the British judicial system; examining its efficacy and whether these steps are performing as they were intended to when incepted into law. There has been some research into both identity parade procedures (Lindsay & Wells, 1985; Gronlund, 2004; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999; Levi, 1998; Wells, 1993; 1998; Steblay *et al.,* 2001; 2010; McQuiston-Surrett *et al.,* 2006; Memon & Gabbert, 2003; Mecklenburg, 2006; Steblay, 2006; 2010; Malpass & Ross, 2007) and the use of Special Measures for vulnerable and/or intimidated witnesses (Swim *et al.,* 1993; Goodman *et al.,* 1998; Landstrom *et al.,* 2005; Davies & Noon, 1991; Nisbett & Ross, 1980), yet there remain a significant number of questions prompting furious debate in the research community. This thesis aims to address many of the contentious issues and satisfy the vital need for empirical data in the contemporary approach to eyewitness identification and testimony presentation methods.

Chapter 2 assesses the effectiveness of the sequential, video based identity parade system recently adopted by many British police forces; V.I.P.E.R. While this move away from the more traditional, simultaneous procedure has been taken in an effort to reduce the enormous costs of organising identity parades, it has also been driven by extensive research purporting a sequential procedure as a means of reducing mistaken identifications and wrongful convictions (McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999; Lindsay & Wells, 1985; Kassin *et al.,* 2001; Kneller *et al.,* 2001). Yet V.I.P.E.R. is very different to the strict sequential procedure employed in much of

the literature, requiring a double presentation of the parade before an identification can be made as well as permitting repeated viewing of parade members (The Police and Criminal Evidence Act, 1984, Code D, 2008). This represents a significant caveat to the extant literature (Steblay *et al.,* 2001; 2010; McQuiston-Surrett *et al.,* 2006; Lindsay & Pozzulo, 1999; Levi & Lindsay, 2001; Levi, 1998) which unfortunately means that the V.I.P.E.R. identity parade format is relatively untested apart from one study (Valentine, Darling & Memon, 2007) which is reviewed later in the thesis. Furthermore, the literature recognises that the exact conditions under which sequential superiority may occur remains unanswered (Wells 1998; 2008; McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001) as does the theory regarding the driving force behind such an effect (Meissner *et al.,* 2005; Gronlund, 2004; McQuiston-Surrett *et al.,* 2006). Chapter 2 aims to resolve the uncertainty that surrounds the V.I.P.E.R. identity parade procedure by testing its effectiveness within the constraints of the British legal system.

Chapter 3 deals with the rising use of Special Measures as a testimony presentation method for adult witnesses who are recognised to be vulnerable or at risk of intimidation (The Vulnerable Witnesses (Scotland) Act 2003, Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003; Hamlyn *et al.,* 2004; Richards, Morris & Richards, 2008). While the use of Special Measures has been widely permitted for child witnesses for a number of years, it is only since 2006 that their criterion for use has been extended to adult witnesses (The Vulnerable Witnesses (Scotland) Act 2003; Richards, Morris & Richards, 2008). Despite the fact that child witnesses have routinely used Special Measures for years, there is little literature which has examined their effects upon juror perceptions and trial verdict decisions. Of the few that have ventured into this area (Landstrom *et al.,* 2005; Goodman *et al.,* 1998; 2004; Swim *et al.,* 1993;, Davies & Noon, 1991) there are small, yet significant indications that the use of Special Measures limits the value of a witness's testimony in the eyes of the jurors; such a witness is perceived as less credible and believable than the same witness who testifies in an open court with no Special Measures. However, we cannot simply extrapolate data from child studies to adult witness studies. The only study to date which has used adult witnesses, who were not vulnerable, is the 2005 paper by Landstrom *et al.*

which corroborates the findings from the earlier child studies but failed to investigate the effect upon trial verdict decisions. Thus, there is a clear need for empirical investigation of the effect the use of Special Measures for adult vulnerable witnesses has upon jurors' perceptions of the witness, the defendant and the trial verdict.

This new pathway through the justice system, from identification to testimony, must be unbiased from all possible influences to ensure a fair and just outcome for all interested parties. This thesis proposes to examine the witness's route through the judicial system and to make a determination as to its fulfilment of its intended goals.

Chapter 2:
Eyewitness Identification

Introduction

If you ask any member of the public to describe an identity parade it is likely that they would describe the enduring image of viewing a group of similar individuals from behind mirrored glass. However this is no longer wholly accurate. In 1996 West Yorkshire Police devised a new method for performing identity parades. This system is called Video Identity Parade, Electronic Recording, or V.I.P.E.R. and has been implemented by a large number of police forces nationally and internationally, replacing the traditional procedure. While the traditional format exhibits the parade members to the eye-witness live at a police station, V.I.P.E.R. allows the police to compile a video-based identity parade which can be administered in the station, at a hospital or the witness' home. A V.I.P.E.R. identity parade is similar in many aspects to the traditional format; between 8 and 12 similar looking individuals are recorded and later shown to witnesses. Each individual is recorded in a twenty second clip which begins with the individual looking directly into the camera and then rotating their head 90 degrees to show their full left profile and then rotating 180 degrees to the full right profile before returning to the initial full face view.

Yet, it is important to acknowledge that researchers in this area have not routinely employed live, traditional, simultaneous parades as part of their methodology. To do so would require an enormous degree of co-ordination, flexibility and would likely have a significant financial implication as well. Instead, researchers have tended to employ the photo line-up, as widely used by police agencies in the United States and Australia. The photo spread is administered in a similar fashion to the live parades; simultaneously. So the witness is presented with an array of similar looking faces on sheet of paper or a computer display. The task is inherently identical to the live, traditional parade in that the faces are still presented to the witness simultaneously; in this study, the photo line-up is employed as best alternative to a live, traditional identity parade.

It may also be prudent at this point to specify the terminology used in relation to identity parade research. A target present, or culprit present, parade is an identity parade constructed with an image of the responsible individual contained within it. Conversely, a

target or culprit absent identity parade is constructed without any image of the responsible person within it. Target absent parades are comprised entirely of 'foils'; people who are selected to fill out the parade because they bear some similarity to the 'suspect'.

There are three outcomes associated with target present identity parades. The first is a 'hit', or more commonly, a correct identification, whereby the witness identifies the culprit/suspect. Secondly, the witness may instead make a mistaken identification of a known foil and wrongly identify one of the foils as the suspect. Thirdly, the witness may choose to reject the entire parade because they do not believe that the suspect is present within that set of individuals.

When a witness views a target absent identity parade there are two potential outcomes. They may elect to reject the entire parade, correctly, as the suspect is not present within it; this is commonly referred to as a 'correct rejection'. The second outcome is that the witness mistakenly identifies one of the foils as the suspect and this is referred to as a 'false identification'.

The simultaneous identity parade format has been shown to encourage the witness to employ a *relative* judgment style (Levi & Lindsay, 2001; Memon & Gabbert, 2003; Lindsay & Pozzulo, 1999; Lindsay & Wells, 1985; Ebbesen & Flowe, 2002; Steblay, Dysart, Fulero & Lindsay, 2001) due to its simultaneous nature. This type of judgement style requires the witness to compare their memory of the suspect to each parade member whilst also comparing the parade members to each other (Steblay *et al.*, 2001; Memon & Gabbert, 2003; Lindsay & Pozzulo, 1999; Kneller, Memon & Stevenage, 2001). Thus, the witness frequently selects the individual who most closely resembles their memory of the perpetrator and is a superior match than any other parade member. This is advantageous when the actual suspect is present in the identity parade and this is clearly supported by the data; the rate of correct identifications elicited using the simultaneous format is estimated to be 75% (Wells, 2008) but has been claimed to be as high as 85-90% (Valentine, 2006).

But if the simultaneous identity parade format performs so highly why has it become necessary to switch to a new, sequential format? It has been widely reported that over half of simultaneous parades organised by police services are cancelled without being viewed (Steblay *et al.*, 2001; Slater, 1995). This leads to huge costs for the judicial system in wasted man-hours and payments for the attendance of foils. The primary reason for such a large cancellation rate is simple: witness anxiety about the possibility of coming face-to-face with the suspect outside of the identity parade. While the police endeavour to minimise contact before and after the identity parade it would be impossible to eliminate any possibility of contact, and thus witness stress, completely. Yet there are additional factors which exert an influence upon cancellation rates. When organising a traditional identity parade it was extremely difficult for the police officer in charge to locate and contact appropriately similar foils, particularly as they were constrained to locally available individuals (Kemp, Pike & Brace, 2001; Valentine and Heaton, 1999). In many cases identifying appropriately similar foils often became a duty for officers on the beat; they would be given a written description of the suspect and then would have to match individuals on the street as possible suitors (Valentine, 2006). Of course not every individual approached would be willing to participate, even for a fee, and, of those who would participate, the organising officer would then have to contact and determine a suitable time for all foils and suspect to attend. Due to such constraints it is not surprising that over 50% of all eye-witnesses who viewed traditional identity parades felt that the parade members were of poor similarity (Slater, 1995). Furthermore, the amount of work required to organise even one identity parade is clearly tremendously expensive and time-consuming and there being a 50% chance of every identity parade being cancelled traditional identity parades are overtly linked with a huge drain of judicial resources (Slater, 1995). Therefore, in a concerted effort to minimise cancellations West Yorkshire Police devised V.I.P.E.R.

However, there are other factors which may influence the eye-witness's failure to reject target absent parades. The first is simply recall delay between exposure to the incident and attending an identity parade. In the United Kingdom traditional identity parades typically require a minimum of ten weeks to organise (Office for National Statistics, 2007). This is

often compounded by the variable delay between the commission of a crime and witnesses coming forward to the police. Research has categorically shown that memory degrades rapidly almost immediately after the event (Pozzulo, Crescini & Panton, 2008; Deffenbacher, Bornstein, Penrod & McGorty, 2008; Ebbesen & Rienick, 1998) and continues deteriorating with increasing time after the event. Coupled with the verbalisation effect (Memon & Bartlett, 2002); which claims that providing police with a verbal description of the suspect makes identification at a parade more unlikely; it is understandable that many witnesses memory of the suspect becomes fragmented quickly after acquisition which make recognition at an identity parade more difficult.

Such a suggestion should explain why the rate of correct rejections is high, not low, yet there is also documented evidence which suggest witnesses rationalise their attendance as suggestive that the suspect is present in the identity parade (Wells, 1993) when he may not appear. Many eye-witnesses will feel that they would have only been called to an identity parade to confirm to the police that they have already located the right suspect; perhaps rationalising 'Why would they call me if they hadn't already arrested the perpetrator? Surely it would be a waste of time if he *isn't* present?'

All of these factors play a role in facilitating the high rate of mis-identifications, and wrongful imprisonment, associated with traditional simultaneous identity parades. In an effort to counter these influencing factors Lindsay & Wells proposed the sequential identity parade procedure in 1985.

While it is true that the primary driving force for the switch from traditional to V.I.P.E.R. was economic factors (Slater, 1995; Steblay *et al.*, 2001; Valentine, 2006) it is widely reported in the literature that the simultaneous identity parade format performs poorly when eye-witnesses are shown an identity parade in which the suspect is *not* present (Pike, Kemp, Brace, Allen & Rowlands, 2000; Lindsay & Wells, 1985; Steblay *et al.*, 2001; Memon & Gabbert, 2003; Lindsay & Pozzulo, 1999; Wells, 1984). It is hypothesised that the reason for this poor performance lies in the judgment style which makes traditional parades so

successful for suspect present parades. While the use of a relative judgment style is advantageous when the suspect is present, because it leads the witness to select the 'closest match' to their memory, it is easily apparent that relative judgments encourage the witness to select the parade member who most closely resembles their memory of the suspect, even though that foil may not be closely similar (Lindsay & Wells, 1985; Wells, 1993; Steblay *et al.*, 2001; Slater, 1995). The rate of false identifications; selecting an innocent individual as the suspect, elicited when employing the traditional, simultaneous identity parade format has been estimated between 20 and 70 percent (Levi, 1998; Wells, Small, Penrod, Malpass, Fulero & Brimacombe, 1998; Memon & Bartlett, 2002; Kemp, Towell & Pike, 1997) with 30% typically cited as common. At the time of writing, over 75% of the 239 wrongful imprisonment cases investigated and over-turned by The Innocence Project in the United States have been due to mis-identification by witnesses (The Innocence Project, June 2009).

In contrast to the traditional format, V.I.P.E.R. employs a partial sequential procedure whereby each individual in the identity parade is presented in turn to the witness rather than simultaneously. Such sequential presentation favours an *absolute* judgment style as it prevents the witness from comparing the individual parade members to each other in order to identify the 'best' match (Memon & Gabbert, 2003; Steblay *et al.*, 2001; Lindsay & Wells, 1985; Lindsay & Pozzulo, 1999; McQuiston-Surrett, Malpass & Tredoux, 2006; Kneller *et al.*, 2001). It encourages the eye-witness to compare their memory of the suspect to each individual parade member in turn and, in a true sequential format, the witness will only view each individual *once*. Thus, the witness must make an identification decision, match or reject, after viewing each face.

It is however imperative to note that there are clear and strident differences between the wholly sequential procedure advocated by the research and the sequential procedure of the V.I.P.E.R identity parade format. During a V.I.P.E.R. parade, the witness may not be informed prior to viewing the parade how many individuals they will see, they are explicitly told not to make their identification decision until the entire parade has been viewed twice

and, perhaps most significantly, after the parade has been viewed twice, the witness may request to view one or more individuals again before they make a decision (Police and Criminal Evidence Act, 1987, 2003). While the strict sequential procedure is advocated by the literature, the researchers fail to agree on a standard, strictly sequential format. There are various proposed models for a sequential identity parade format; the majority of studies advocate, and have employed, a wholly sequential format whereby the witness must give a final decision (select or reject) to the currently presented individual before advancing to view the next (Lindsay & Wells, 1985; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999; McQuiston-Surrett *et al.*, 2006; Steblay *et al.*, 2001; Valentine *et al*., 2007; Wells, 1998). There are also other factors which vary between proposed sequential parade procedures; the question of single- versus double-blind presentation (Wells, 1993; 1998; Levi & Lindsay, 2001; McQuiston-Surrett *et al.*, 2006; Steblay *et al.*, 2001; Lindsay & Pozzulo, 1999), whether the participants should be told how many individuals they will see (Lindsay & Wells, 1985; Levi & Lindsay, 2001; Steblay *et al.*, 2001; McQuiston-Surrett *et al*., 2006; Lindsay & Pozzulo, 1999; Wells, 1993; 1998; 2001; Memon & Gabbert, 2003; Valentine *et al.*, 2007; Kassin *et al.*, 2001; Valentine, 2006; Valentine & Heaton, 1999) and the use of sequential photo parades to assess sequential superiority rather than moving images or video of the parade members (Steblay *et al.*, 2001; McQuiston-Surrett *et al.*, 2006; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999; Valentine *et al*, 2007; Memon & Gabbert, 2003; Wells, 1998; 2001).

From the published literature there does appear to be significant benefits associated with such a presentation style. Of paramount importance is the claim that the sequential presentation method significantly reduces the number of false identifications of an innocent foil during target absent identity parades. Within the forensic community it has been widely accepted that sequential presentation of an identity parade maintains the high hit rate associated with simultaneous parades for target present situations whilst significantly lowering the number of mis-identifications (Steblay *et al.*, 2001; 2011; Lindsay & Wells, 1985; McQuiston-Surrett *et al.*, 2006; Lindsay & Pozzulo, 1999; Valentine, Darling & Memon, 2007). As discussed earlier though, many of the claims of sequential superiority stems from

research which has not used the V.I.P.E.R. format employed in the United Kingdom; instead many studies have simply used sequential presentation of still photographs. V.I.P.E.R. is a wholly different format which utilises moving images of each parade member presented sequentially. At this time there has been only one study which has examine this aspect of the V.I.P.E.R. identity parade system; Valentine, Darling & Memon's 2007 paper which tested the effect of both moving images and strict instructions on the rate of successful outcomes for target present and target absent identity parades.

Valentine, Darling & Memon's 2007 paper is, to date, the only published study which has employed a methodology which truly reflects the current practice for all V.I.P.E.R. parades conducted in the United Kingdom. The objective of this study was two-fold; firstly to assess the effect of using moving video images as parade stimuli over static full-face images and secondly whether the current instructions for V.I.P.E.R. parades are adequate to inhibit relative judgments or whether strict viewing guidelines would elicit a significant improvement in mis-identifications.

The design of this study had three between-participants factors, each with two levels; procedure (existing versus strict), image format (moving versus static) and culprit presence (present versus absent). 223 student participants from the University of Aberdeen viewed a live, staged crime before being asked to return at a later date to view an identity parade. 14 participants withdrew from the experiment for various reasons and so data was drawn from the remaining 202 participants. Upon returning for the identity parade, the participants were randomly allocated to one of the 8 conditions. All parades were introduced using the standard guidance and instructions given to witnesses before a parade is begun, stressing that if the witness feels that they cannot make an identification they must say so and not make a guess. All parades were administered double-blind in that neither the witness of the experimenter knew if the culprit was present in the parade. Static photo line-ups were administered via a 15 inch computer monitor. In the 'existing' condition, the V.I.P.E.R. images were played in their entirety, twice, before allowing an identification to be made and furthermore, allowing witnesses the opportunity to view any image they wished to see again,

for however long they desired. This methodology follows the current PACE Code D instructions. In contrast, the 'strict' condition presented the video images of the parade members one at a time and required the witness to make one of three decisions: 'yes' – indicating a positive identification which immediately ended the identity parade, 'no' – indicating a rejection and prompting the presentation of the next image, or a request to view the same individual for a second time. Once a parade member has been rejected, the participant was not permitted to view it any further. These strict instructions are in line with much of the research in this area (Cutler & Penrod, 1988; Lindsay, *et al.*, 1991; Lindsay & Wells, 1985; Steblay *et al.*, 2001; Kemp *et al.*, 2001; Levi & Lindsay, 2001; McQuiston-Surrett *et al.*, 2006).

The authors hypothesised that the 'strict' condition would yield fewer mistaken identifications in the culprit absent parades compared to the existing procedure, as the strict instructions make it much more difficult to make relative judgments. Due to this, it was also hypothesised that the strict condition would generate fewer correct identifications in the culprit present parades. The second hypothesis was that the moving images would be associated with a higher rate of correct identifications based upon the encoding specificity principle (Tulving & Thomson, 1973).

Hierarchical log-linear analysis of the culprit present data showed that while there was a significant effect of instructions, whereby the strict instructions provoked significantly fewer identifications, correct identifications of the culprit and mistaken identifications. The analysis revealed there to be no significant difference in the rate of identifications between the moving and static image formats; thus the predictions of the encoding specificity principle were not valid, as there was no increase in the rate of correct identifications. However, the culprit absent data revealed there to be no significant effect of instructions on the rate of correct rejections or misidentification but did show a statistically significant main effect of image format. The proportion of correct rejections when viewing the moving images was significantly greater than that obtained when witnesses viewed the static photo line-ups; although the size of this effect was small ($\Phi = 0.24$).

Based upon their results, the authors recommended that the existing procedure currently in force across the United Kingdom be preferred over the strict procedure. This is because the strict procedure was associated with a significant decline in parade sensitivity; a reduction in the number of correct identifications in the culprit present parades. This recommendation was further reinforced by the evidence suggesting a small, yet significant, benefit of moving images for culprit absent parades, whereby the moving images elicited fewer mistaken identifications than the static photo images.

While many authors link the advantageous decrease in mis-identifications purely with the specific judgment style fostered by sequential presentation there are many other possible factors which may be responsible for such an effect. The delay between incident exposure and recall is much shorter when V.I.P.E.R. parades are used; dropping to around ten days (although they can be organised within hours) from an average of ten weeks with the traditional format (Office of National Statistics, 2007). As a result, although much of memory is forgotten rapidly after encoding without active remembering, the eye-witness' memory of the suspect should be stronger than it typically would be after ten weeks (Memon, Hope & Bull, 2003; Deffenbacher, Bornstein, McGorty & Penrod, 2008). Furthermore, when individuals are recorded and entered onto the V.I.P.E.R. database their status is noted; whether they are innocent volunteers, suspected of criminal activity or convicts (West Yorkshire Police, 2009). This is information that would not be ordinarily available to organising police officers during traditional 'live' parades and therefore is a possible route to increasing false identifications during target absent parades. When using the V.I.P.E.R. format, even if the eye-witness selects an innocent foil as the suspect the organising V.I.P.E.R. officer will check the status of such an individual and a false identification is unlikely to proceed any further.

There has recently been a further argument that sequential presentation of foils does not elicit a change in judgement style but simply encourages the witness to adopt more conservative criteria for recognition (Ebbesen & Flowe, 2002; Meissner, Tredoux, Parker & MacLin, 2005). This hypothesis is based on signal detection theory and claims that the eye-

witness may not be choosing to use an absolute judgement over a relative judgement but rather that they increase the criterion required to make a positive identification of the suspect. Much of the evidence supporting the claim that witnesses use an absolute judgement style when viewing sequential identity parades has derived from self-report measures (McQuiston-Surrett *et al.*, 2006; Meissner *et al.*, 2005) and thus should be regarded cautiously. In addition, sequential identity parades are routinely compiled from a national database of images, therefore the quality of the foils is significantly improved (Valentine & Heaton, 1999; Steblay *et al.*, 2001) and perhaps a more robust criterion is employed by the witness to distinguish and identify the perpetrator (Ebbesen & Flowe, 2002; Meissner *et al.*, 2005). Gronlund (2004) tested the efficacy of both the high criterion match theory and the decision strategies theory and found strong evidence to support the change in decision strategies as the explanation for what occurs when eye-witnesses view identity parades in place of signal detection theory.

There are also claims that eye-witness identification is aided through the use of video identity parades as the movements made by the parade members are more natural and similar to those made during encoding of the suspect's image as the crime is committed. Tulving & Thompson (1973) stated that 'viewing a moving image at recall increases the likelihood of recognition compared to still image'. This is because the eye-witness will have likely seen the suspect moving around throughout the crime and the encoding specificity principle claims that recall, and recognition, will be facilitated if the stimulus is presented in the same manner in which it was originally encoded (Tulving & Thompson, 1973). The results from Valentine, Memon & Darling's 2007 study found a small but significant increase in correct rejections for target absent parades when the eye-witnesses viewed a moving CCTV video and a moving identity parade. It has been suggested that the moving stimulus gives the witness a wider range of cues or perhaps reveals a facial distinctiveness which later aids identification and encourages witnesses to be more stringent in their recognition criteria. Despite this, there remains a body of evidence which reports no significant benefit to eye-witness identification when moving images are used over still photos (Henderson,

Bruce & Burton, 2001; Bruce, Henderson, Newman & Burton, 2001; Hancock, Bruce & Burton, 2000).

A further theory which may provide some answers as to why sequential parades elicit a lower rate of false identifications is the three-quarter face effect (Bruce, Valentine & Baddeley, 1987). The effect is claimed to have a small but significant advantage for recognition when presented to witnesses (Liu & Chaudhuri, 2001; Hancock, Bruce & Burton, 2000), particularly when it is also seen alongside the full face view. Clearly, the crime video provides both full face and 3/4 views for the witness to assimilate and then at the recognition stage the V.I.P.E.R. segments provided a full range of views to the witness; from full face through to each profile, which lends itself favourably to benefiting from the 3/4 effect for recognition.

However, it should be noted that the adoption of sequential procedures has not been without its controversies. From the beginning the proponents of sequential presentation called firmly for strict conditions under which sequential identity parades should be run. These included a double-blind procedure and true sequential testing; that the witness *must* reject the individual before them as the suspect before viewing the next image; that the witness cannot change their rejection/acceptance decision and that they may not know how many images they will view (Lindsay & Wells, 1985; Wells, 1993; Levi & Lindsay, 2001; Valentine *et al.*, 2007; Memon & Gabbert, 2003). However, these conditions have not been met by worldwide police forces; it is often impossible to run a double blind procedure and true sequential testing is contrary to the provisions of the Police and Criminal Evidence Act Code D, which provides for all eye-witnesses to view the entire video parade *twice* before making a decision (Home Office, 2004). Clearly then the current UK procedure for sequential identity parades is not truly sequential and can reasonable be claimed to have a simultaneous component.

Despite this issue the introduction of the V.I.P.E.R. system has been deemed a success by the 34 UK police forces which have implemented its use. All forces now operate several

operations for recording images and have a dedicated team of V.I.P.E.R. officers. All images recorded are contributed to the central UK database which can be accessed and utilised by any member force to compile identity parades. This has helped to reduce man hours spent in seeking suitable foils locally and organising attendance of suspect/foils and witness simultaneously. Most importantly, because there is no longer any possibility of witness and suspect inadvertently meeting, the number of V.I.P.E.R. parades cancelled is just 5% (Slater, 1995; Home Office, 2003).

Thus, it has been widely accepted that the traditional identity parade format is equally as good, if not superior, to the V.I.P.E.R. format when the suspect is present whilst the V.I.P.E.R. format is significantly superior for target absent identity parades (Lindsay & Wells, 1985; Memon & Gabbert, 2003; Valentine *et al.*, 2007; Lindsay & Pozzulo, 1999; Steblay *et al.*, 2001; 2011; Wells 1993; Levi & Lindsay, 2001). Yet despite 81% of researchers in the community advocating the superiority of the sequential method for identity parades (Kassin, Tubb, Hosch & Memon, 2001; McQuiston-Surrett *et al.*, 2005; Valentine *et al.*, 2007) there remain questions regarding the precise conditions under which sequential presentation is superior to simultaneous. Much of the available literature has tested V.I.P.E.R.'s performance under a wide range of conditions from witness stress and arousal (Pozzulo, Crescini & Panton, 2008) to change of suspect appearance (Henderson, Bruce & Burton, 2001), operating procedures (Wells, 1993) and judgement styles (Memon & Gabbert, 2003; Lindsay & Wells, 1985), and while the data has provided valuable information about video identity parades there has not yet been a clear undertaking to examine whether the sequential superiority effect occurs under the simplest conditions; recall delay and CCTV crime video colour.

CCTV video colour was introduced as a factor due to recent advances in CCTV image quality. Typically, black and white, or greyscale, CCTV cameras were much more sensitive to light and offered greater resolution than colour camera, all for a much cheaper cost. However, in recent years, the differences in cost have been greatly reduced and colour CCTV cameras now offer the same level of resolution as greyscale cameras. Although the

greyscale CCTV cameras do still offer superior light sensitivity, this superiority is often marginal and there are some colour CCTV cameras can capture images in light levels as low as 0.3 lux, although these are associated with greater costs. In spite of these costs, it is believed that colour CCTV cameras yield better recognition capability than greyscale CCTV cameras (http://www.cctvsentry.com/faq.htm), allowing CCTV operators to specify clothing colour to police searching for a particular individual seen on camera (Davies & Thasen, 2000).

Research in this area has not supported such a claim though. Laughery, Alexander & Lane (1971) found there to be no advantage in recognition of mug shots when they were presented in colour rather than greyscale; a finding replicated by Wogalter & Laughery (1987). Similar negative results have been reported by both Bruce *et al.*, (1999) and Shepherd, Ellis & Davies (1982) when the task was to identify one target face contained within a photographic array. Davies & Thasen (2000) conducted an experiment where 80 participants were instructed to watch a surveillance video, shown in either colour or monochrome, and note any suspicious events in a log. After checking that all participants had recorded the event of interest, the participants were asked to describe the individual they saw and to attempt to identify them from a photographic array, again presented either in colour or monochrome. Their results showed that, surprisingly, the participants made little use of the colour cues in their descriptions of the clothing worn by the target; 40% described the colour of her coat but only 17% described the colour of her trousers. Just 15% of the participants successfully identified the target individual from the photo array while 60% made a false identification.

The authors found evidence for a main effect of video colour; those participants who viewed the colour CCTV tape were less caution in their study of the photographic array and were more likely to choose compared to those participants who viewed the monochrome CCTV tape. Those who viewed the monochrome CCTV tape were superior on just two points; firstly, they provided more detailed descriptions of the target's hairstyle and more

importantly, they made fewer false identifications than those who viewed the colour CCTV tape.

On the basis of these results, the author claimed that colour has an inconsistent and unpredictable effect upon eyewitness testimony. While those who viewed the colour CCTV tape gave more spontaneous descriptions of the target's clothing, they were also more likely to wrongly identify an innocent foil. Davies & Thasen suggest that the witness' confidence in their own ability to make an identification was boosted through the presence of colour and that this in turn primed them with a greater readiness to respond, regardless as to the presence or absence of the target in the photographic array. Due to this, while the authors recognise that colour may be a valuable factor in improving person identification, they recommend that CCTV operators be trained appropriately to recognise the advantages and potential disadvantages of using colour cues for identification purposes.

There appears to be only one study which suggests that colour may have a beneficial effect upon face identification which was published in 1985 by Tickner & Poulton. Their study asked their volunteers to observe surveillance footage for either 1, 3 or 4 hours and then asked them to search for 3 or 12 target faces within a photographic array. Within their experiment the presence or absence of colour and the faces in the photo array were systematically varied. Their results are interesting though, particularly in terms of the V.I.P.E.R. procedure; Tickner & Poulton found a small, yet significant advantage for colour in terms of identification. Those participants who viewed the colour CCTV surveillance footage were more likely to make correct identifications of the targets, regardless of whether the photographic array was presented in colour or greyscale.

In light of these conflicting results, it was decided to examine the effect of CCTV colour as an additional factor which may affect eyewitness identification, in an effort to fully resolve the issue.

Furthermore, the present study was designed to adhere to the current procedure used by all UK police forces laid down in The Police and Criminal Evidence Act Code D (2008, Home Office). Therefore the 'eye-witnesses' throughout this experiment were required to view the V.I.P.E.R. identity parade *twice*, divergent from the standards advocated by the research community (Lindsay & Wells, 1985; Valentine *et al.*, 2007; Lindsay & Pozzulo, 1999; Wells, 1993; Levi & Lindsay, 2001).

Therefore, it was hypothesised that there would be a significant difference in correct identifications of the suspect between the simultaneous and V.I.P.E.R. identity parade format for target present parades and that V.I.P.E.R. parades would elicit significantly fewer false identifications than the simultaneous parade for target absent parades. In addition, drawing from the literature (Deffenbacher, Bornstein, Penrod & McGorty, 2008; Ebbesen & Rienick, 1998), it was further expected that both parade types would suffer from lower rates of success (less hits and more mis-identifications) in the longer recall delay conditions and that those eye-witness who viewed the black and white CCTV crime video would perform poorer than those who had viewed the colour CCTV crime video.

Method

Design

The experiment was conducted in two parts. Firstly, the participants viewed a CCTV video of a simulated crime before returning at a later time to take part in an identity parade. The experiment had four between-subjects factors, each of which had two levels: *Parade Type* (simultaneous versus V.I.P.E.R.), *Recall Delay* (2 hours versus 24 hours), *CCTV video* (Black & White versus Colour) and *Target Presence* (target present versus target absence).

The dependent variables were different according to the presence or absence of the target in the identity parade. The dependent variable for the target absent parades was the number of correct rejections of the entire identity parade. However, the dependent

measures for the target present parade were the number of correct identifications *and* the number of mistaken identifications.

Participants

320 participants were recruited from the population of the University of Glasgow. Many of the participants were level one Psychology undergraduates who participated in exchange for course credit. The average age of the participants was 24.48 years (SD = 9.64, range: 17-67 years). Of the 320 participants 185 were female. The participants were randomly allocated to one of 8 conditions resulting in 40 participants in each group. These 40 participants in each group were further split as 20 viewed a Target Absent parade while the remaining 20 viewed a Target Present parade (see Figure 1).

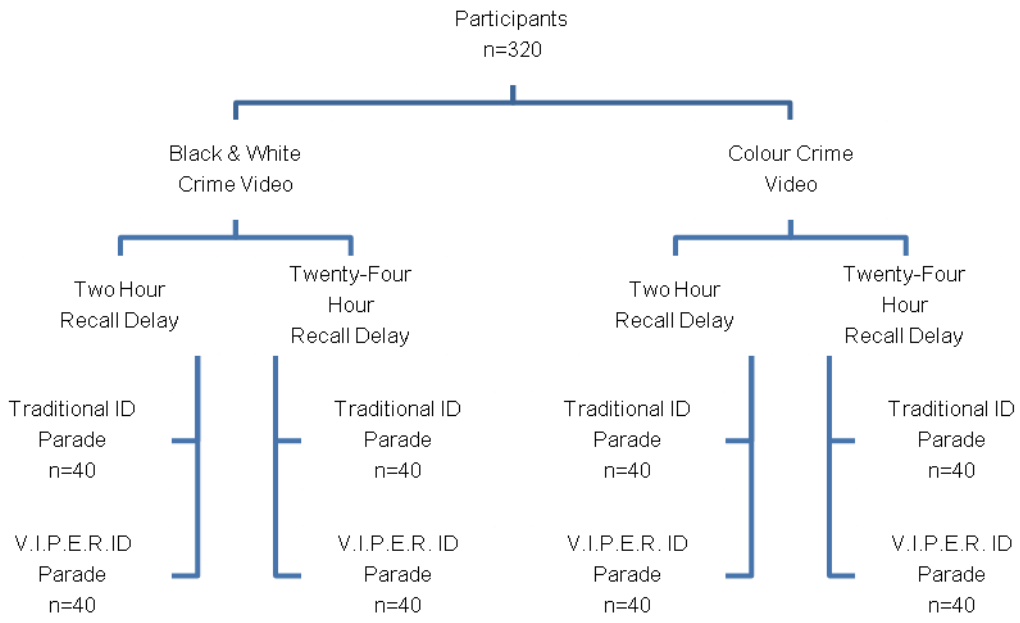


Figure 1: Breakdown of the configuration of the experiment series.

Upon completion of the experiment each participant was fully de-briefed, thanked and received payment for their time; either course credit or £3.

Stimuli

The crime video watched by all participants was filmed in a department waiting room in the University of Glasgow. In order to avoid any negative effects upon the 'witnesses' a low impact crime (theft) was selected to be the subject of the 'crime' video. Three staff members were recruited to enact the 'crime'; theft of a wallet. In order to retain access to a broad subject pool the three staff members were selected based on how little contact they had with the Psychology undergraduates. Thus the participants would have no pre-existing familiarity with the 'criminal' or his 'victim'. The three staff members were given a detailed overview of the aims of the experiment and the conditions under which the video was to be filmed. There were to be three roles within the context of the video: criminal, victim and experimenter who called away the victim. The video would show the victim arriving in the waiting room followed by the criminal. After a few seconds an experimenter would call the victim, who leaves their backpack in the waiting room. Whilst alone, the criminal rummages through the backpack and steals a wallet before casually leaving.

The stimulus video was designed to mimic a CCTV video therefore a video camera was positioned relatively high on a wall for the recording, approximately 8 feet from the ground and 10 feet across the room from the incident. The video was recorded after the department had closed in order to eliminate any extraneous faces from the stimulus. After some last minute rehearsals and alterations the CCTV video was recorded three times, once with each staff member playing the role of the criminal, victim or experimenter. The video was recorded in both a colour and black and white version and was approximately 80 seconds in duration. The video clearly shows a full range of facial views of the suspect, from full profile through to full face. The resolution of the black and white CCTV video was identical to the resolution of the colour CCTV video as no changes were made with the exception of reducing the saturation level to remove all colour cues. No other changes were made to the stimulus crime videos as a control measure.

After recording was completed the three versions of the stimulus video was shown to a panel of four staff members in order to identify the most plausible 'criminal'. Once selected

some minor editing was performed. All sound was removed from the clip and the beginning and end shaved to remove unwanted images. A still image of the criminal was then taken from the video and shown to ten random individuals who were asked to describe the individual. These descriptions were recorded and used in the next stage of stimulus development.

Still images taken from both colour and black & white versions of the stimulus crime video. 'Suspect' is seated nearest the door.

To compose both the simultaneous and V.I.P.E.R. identity parades the experimenter selected 30 images from a database of approximately 800 faces held by the Department of Psychology's Face Processing Lab. The images were selected on a basis that they matched a written description of the image taken from the video clip of staff member who portrayed the 'criminal', as described by ten random individuals. These 30 faces were then presented to a group of 15 participants who were asked to rank the faces in order of

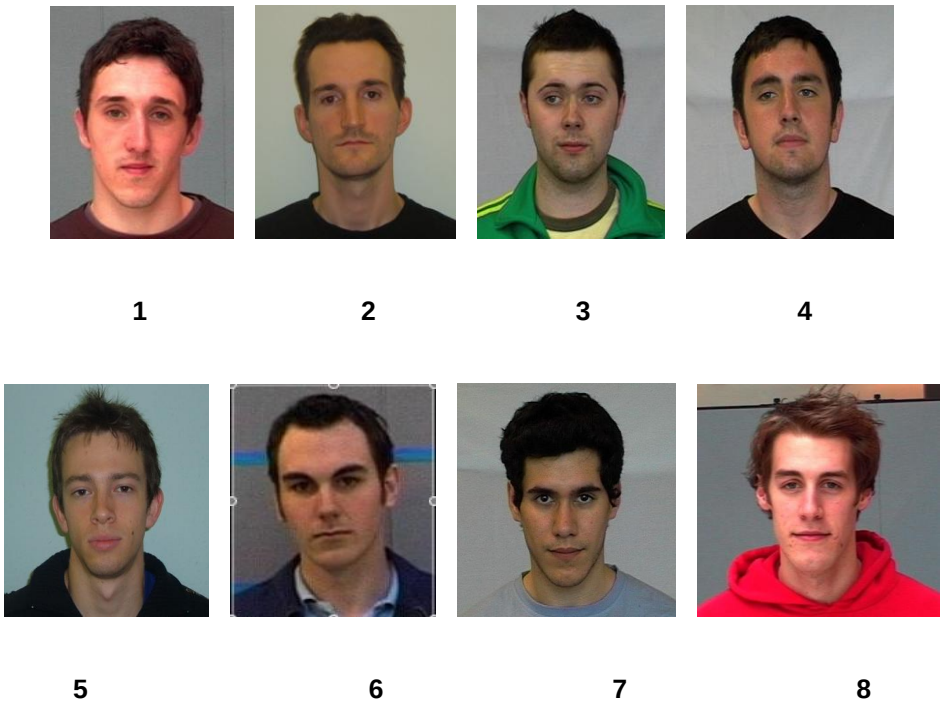
similarity to the target face. The participants were not permitted to discuss their rankings or to collaborate. From the results the 8 highest ranked faces were then selected to compile all four of the required identity parades. In accordance with PACE regulations, the 'criminal' was shown the final parade images and their approval was sought and was given.

Foil selection was carried out in this manner, different to that used in actual V.I.P.E.R. parade composition, because the experimenter did not have access to the national database of V.I.P.E.R. images and could not utilise the selection process of the database. The Face Processing Lab has a relatively large database of faces recorded in typical V.I.P.E.R. presentation which have been gathered from a number of open day events and social mixers attended by potential students from across the United Kingdom.

The eight individuals ranked by participants as most similar to the 'suspect'.

Obviously, there were practical implications which made the use of a live, traditional identity parade impossible as an experimental option for the simultaneous condition. The next best methodology was to use a simultaneous photo line-up, as commonly employed in police practice in the United States and Australia and which has been widely used as a representative simultaneous presentation of an identity parade (Lindsay & Wells, 1985; Levi & Lindsay, 2001; Steblay *et al.*, 2001; 2011; McQuiston-Surrett *et al*., 2006; Memon & Gabbert, 2003; Valentine *et al.*, 2007; Kneller *et al.*, 2001; Clark & Davey, 2005).

The simultaneous identity parades were composed of 8 images; the 7 highest ranked 'foils' and the 'suspect'. Each of these images showed the individual from the chest up with a number positioned below them.



Example of the simultaneous identity parade used in the series of experiments. In this case the suspect is present, in position 2.

It is a common misconception that foils for identity parades are all required to wear similar styles and similar coloured clothing. The foils recruited to fill a parade, whether traditional or V.I.P.E.R., attending wearing the clothing they chose for themselves. Standardisation of clothing across the parade is not thought to exert a significant effect upon witness perceptions (Davies & Thasen, 2000), as it would be unlikely that the culprit would be wearing the same clothes as during the commissioning of the crime. As such, it is not routinely controlled for in applied V.I.P.E.R. parades. The standard V.I.P.E.R. suite set-up dictates recording the moving clips against a light/grey background to provide optimum contrast. As the V.I.P.E.R. database utilises clips on a national basis, it is not unexpected that there would be some variation across the database. On this basis, it was decided that

modifying the backgrounds of the selected video clips would be unnecessary and bring no additional benefit to the study.

The V.I.P.E.R. parades are comprised of video sequences of each individual foil and the suspect. The eye-witness can see only the head, shoulders and chest of each individual as they rotate their head throughout 180 degrees. An identifying number is displayed in the top right corner for each sequence.

Example of one individual's V.I.P.E.R. sequence. Each sequence is approximately 15 seconds duration.

Apparatus

The crime video was shown to each participant on a 17 inch widescreen laptop screen. The viewing of the identity parades took place in a private room with just the experimenter present. Both simultaneous and V.I.P.E.R. parades were displayed on a 21 inch computer monitor.

Participants indicated their decision after viewing on a response sheet (Appendix 1).

Procedure

Participants were seated before a 21 inch widescreen monitor in a quiet room. The experimenter explained that they would be watching live feed from a CCTV camera from a nearby waiting room. The experimenter then left the room and the participant viewed the appropriate 1 minute 25 second 'crime' video. After the video had ended the experimenter re-entered the room and arranged to meet the participants later; 2 hours or 24 hours

depending upon the condition they had been allocated to, at which time they would view the appropriate identity parade.

Upon returning, participants were seated in an office before a 17 inch laptop computer. The experimenter explained that they would now view an identity parade which the suspect may or may not be present. Those 160 participants who viewed the simultaneous identity parade were given standard police instructions (Appendix 2) and permitted as long as required to make a decision. The 160 participants who viewed the V.I.P.E.R. identity parades were read standard police instructions, which explicitly instructed the participant *not* to make a decision before viewing each individual twice and were then left to view the parade. After the V.I.P.E.R. parade had completed, the experimenter returned to the room and asked the participant if they felt able to make an identification. If they expressed themselves to be unsure or a desire to view certain individuals again, in accordance with PACE instructions, the participant was permitted to view the requested images again, to freeze the images and to view them for as long as they felt necessary until they were able to either make an identification or elected to reject the parade.

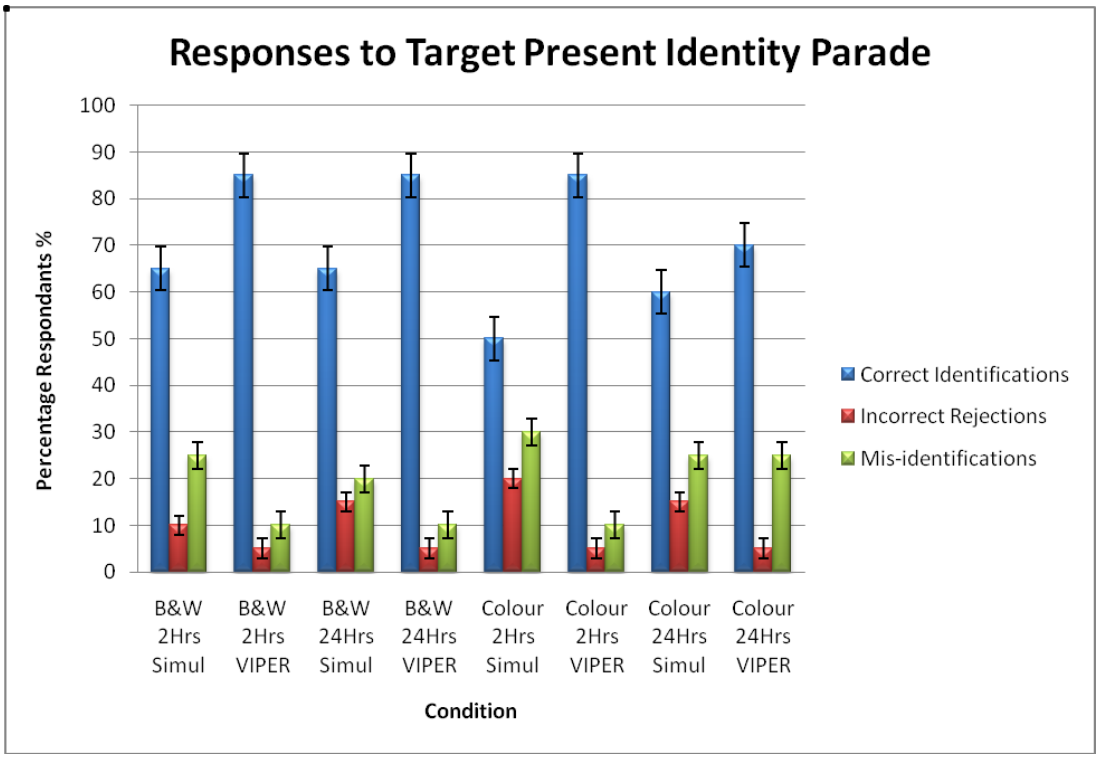
The experimenter stood at the rear of the office, out of the participant's view, throughout this stage of the study. After each participant had made a decision and indicated it upon the response sheet, the experimenter debriefed them, answered any questions and issued payment.

Results

The data were separated prior to analysis into Target Present responses and Target Absent responses. This is because the two methods of identity parade measure two very different constructs; Target Present parades are a test of witness memory whilst Target Absent parades are a measure of witness sensitivity.

Target Present Identity Parades

The following data is generated from the 160 participants who viewed a target present identity parade. The rate of correct identifications elicited during all eight experimental conditions is shown in Graph 1. The data indicate that the simultaneous identity parade format was out-performed by the V.I.P.E.R. parade format under each of the four experimental conditions (see Graph 1).



Graph 1: Percentage correct identifications by each parade style for each of the four experimental conditions. Error bars reflect the standard error (SE).

The overall rate of correct identifications from the 160 participants was 70.6%. Of those who made a correct identification 57.5% viewed the V.I.P.E.R. identity parade; 49.5% viewed the identity parade 24 hours after exposure to the suspect and 53.1% viewed the black & white crime video.

The data was further analysed using nominal logistic regression, as the data are categorical with 3 distinct levels (correct identification, mis-identification and incorrect rejection), to

identify any predictor variable which may exert a significant effect upon the probability of the witness to correctly identify the suspect. The results of this analysis revealed there to be no significant relationship between any predictor variable(s) and the probability of eliciting a correct identification (χ^2 = 45.42, df=62, p>0.05, Φ = 0.73).

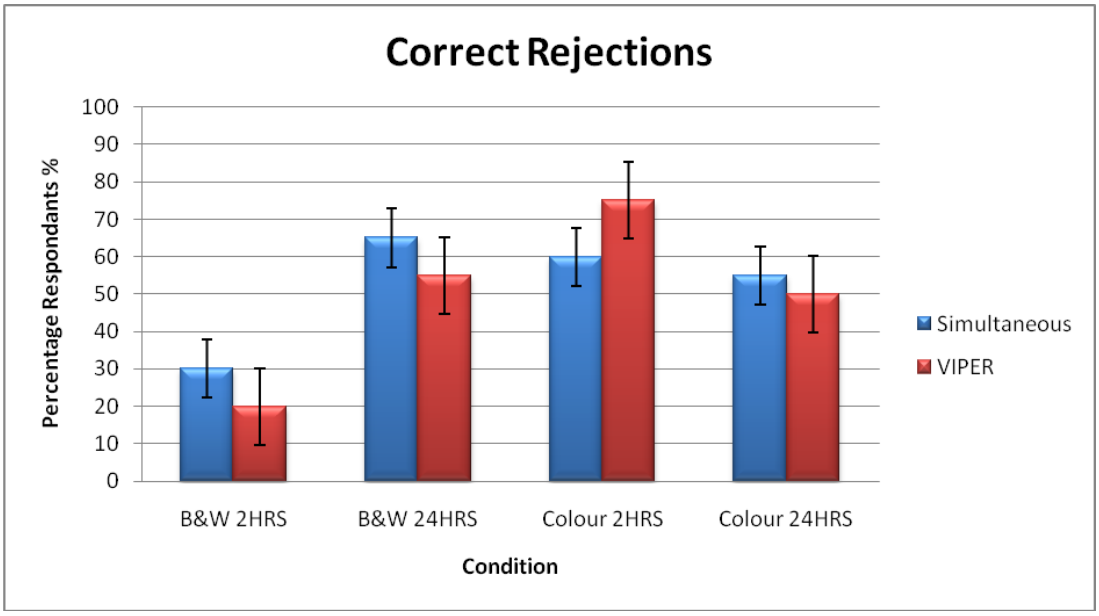
Gender

Gender was included in the logistic regression model as a potential predictor of identity parade outcome however there was no evidence to suggest that witness gender significantly affects the likelihood of making a correct identification (p>0.05).

Target Absent Identity Parades

The following data is generated from the 160 participants who viewed a target absent identity parade. The rate of correct rejections elicited during all eight experimental conditions is shown in Graph 2. The data indicates that the simultaneous parade format was generally superior to the V.I.P.E.R. format in eliciting fewer mis-identifications. The number of correct rejections was particularly poor using both identity parades formats when eye-witnesses viewed the black & white CCTV crime video and returned 2 hours later to view the identity parade. The V.I.P.E.R. format appeared to be superior to the simultaneous format when participants had viewed the colour CCTV crime video and returned to view the identity parade 2 hours later.

The overall rate of correct rejections across all 160 participants was 52.25%. Of those who correctly rejected the entire parade 58.5% viewed the colour crime video; 54.9% viewed the parade 24 hours after exposure to the suspect and 51.2% viewed simultaneous parade.



Graph 2: Percentage of respondants who correctly rejected the entire parade. Error bars reflect the standard error (SE).

As the data were once again categorical and discrete, further analysis using binary logistic regression was undertaken. The resulting model indicated there to be only one predictor variable to exert a significant effect upon the outcome; recall delay ($p<0.05$). The Pearson's test revealed this model to be significant ($\chi^2 = 64.99$, df = 45, $p<0.05$), (odds ratio (OR) = 1.5). All other terms analysed in the model had no significant effect upon the outcome of the parade. Therefore, there was no significant effect of parade type and the null hypothesis cannot be rejected; there is no significant evidence that the V.I.P.E.R. parade format is superior to the simultaneous format for target absent parades.

From the logistic regression analysis, we can state that recall delay is a significant predictor of identity parade outcome and that the increase in recall delay from 2 hours to 24 hours increases the likelihood of making a correct rejection by 1.5 times.

Gender

Gender was included in the regression model as a potential predictor of parade outcome however analysis revealed no effect of gender on parade outcome (p>0.05) and was thus excluded from the regression model.

Discussion

It was anticipated that the current study would yield similar results to those already well established in the research community; that there would be no significant difference in hit rates between the two parade styles for target present parades and that the V.I.P.E.R. format would be significantly superior than the simultaneous format in minimising the rate of mis-identifications (Lindsay & Wells, 1985; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999; Wells, 1993; Memon & Gabbert, 2003; Valentine *et al.*, 2007; Steblay *et al.*, 2001; 2011). However, these anticipated effects did not materialise. The results suggest that there is no significant difference between parade formats for either target present or target absent identity parades.

It has been widely claimed that the sequential identity procedure retains the high hit rate exhibited by simultaneous parades while reducing witness anxiety, delays and costs (Steblay *et al.*, 2001; Lindsay & Wells, 1985; Valentine *et al.*, 2007; Pike *et al.*, 2000), a claim that is supported to some extent here; of the 80 participants who viewed the simultaneous parade 60% of them correctly identified the suspect. However it cannot be ignored that 25% of these 80 participants mis-identified an innocent foil while the remaining 15% incorrectly rejected the entire parade.

There have also been claims that for target present parades the hit rate may be compromised by the sequential procedures' advocacy of an absolute judgment style (Memon & Gabbert, 2003), suggesting that while sequential processing reduces false identifications it may also be limiting the number of correct identifications (Wells, 2008). It is

suggested that this is because of the absolute judgment style which sequential identity parades reportedly fostered in eye-witnesses. Pozzulo & Lindsay (1995) and Memon & Gabbert (2003) assert that while the witness may observe that the image of the suspect is familiar it may be tempting to reject the individual in search of a better criterion match. The current study does not support this claim as 65 of the 80 participants (81.25%) who viewed the sequential parade made a correct identification, 9 (11.25%) incorrectly identified an innocent foil while the remaining 6 participants (7.5%) incorrectly rejected the entire parade. Certainly the datum from the current study appears to suggest that the sequential V.I.P.E.R. format elicits a higher rate of correct identifications compared to the simultaneous parade format and decreases the rate of mis-identification by more than half, as compared to the simultaneous format.

While these results do not support those of Memon & Gabbert's 2003 study it must be noted that the procedure used in Memon & Gabbert's 2003 study was completely sequential; the identity parade was stopped as soon as the witness made a selection, as advocated by proponents of the sequential format. This is a contravention of the legislative guidelines for V.I.P.E.R. identity parades, which requires that sequential parades are shown in their entirety twice before an identification can be made; a stipulation to which the current study adheres. It may be then that the conclusions drawn from this data are more reflective of the actual outcomes observed in real-life identity parades and so it may be that the sequential superiority effect does not manifest in sequential parades which are held in accordance with PACE Code D and the Lord Advocate's Guidelines.

It was anticipated that there would be evidence of a significant sequential superiority of the V.I.P.E.R. format compared to the simultaneous format for target absent parades; however this was demonstrated to not be evident. This is contrary to a large proportion of the published data which clearly state that this pattern should be exhibited (Wells, 2001; Lindsay & Wells, 1985; Lindsay & Pozzulo, 1999; Levi & Lindsay, 2001; Memon & Gabbert, 2003; Valentine *et al.*, 2007; Wells, 1993; Steblay *et al.*, 2001, 2011) with the sequential format performing significantly better than the simultaneous format. Why are the results from the

current study, for both the target present and target absent parades, so divergent from the hypothesised effects which are widely reported in the literature?

It is speculated that the issue at the heart of the findings lies with the procedure employed throughout the entire experiment series. It was decided that the experimental procedure should adhere to current UK police procedures and legislation to enhance the external validity of the experimental series (Bornstein, 1999; Pozzulo, Crescini & Panton, 2008). At this time, PACE Code D stipulates that witnesses attending a V.I.P.E.R. parade be explicitly instructed to refrain from making a decision until they have viewed all the images twice (The Police and Criminal Evidence Act Code D, 2008, Home Office). This is therefore *not* a truly sequential judgment task. This procedure is contrary to the procedure for sequential parades as advocated by experts within the community and it is different on three major points. Firstly it is widely recommended that all sequential identity parade procedures be *strictly* sequential. This is to say that the eye-witness may only view one individual at a time then make a decision. If the witness identifies the individual as the suspect then the parade is immediately ended (Valentine *et al.,* 2007; Memon & Gabbert, 2003; Lindsay & Wells, 1985) while if they reject the individual before them as the suspect the eye-witness proceeds to view the next foil. Secondly, once the witness has made a decision on the individual before them they are not permitted to later change that decision (Valentine *et al.*, 2007; Memon & Gabbert, 2003). The witness may not later identify a previously rejected individual and equally they cannot reject the individual they earlier identified as the suspect. Finally the literature advises that the eye-witness is not told beforehand how many individuals will appear in the identity parade (Valentine *et al.*, 2007; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999). This measure is designed to prevent the witness counting individuals and feeling pressured to make an identification as the parade comes towards its conclusion.

Clearly the stipulations of PACE Code D contravenes the first and third of these recommendations as the witness is required to view the entire V.I.P.E.R. parade twice which also allows the witness to determine how many individuals are shown throughout the whole

parade. It is not unreasonable to propose that in its current format V.I.P.E.R. is not a truly sequential identity parade but is a hybrid state between simultaneous and sequential formats. This model may also account for the exceptionally high level of correct identifications generated by the V.I.P.E.R. format for target present parades: the witness may make a tentative, internal identification on the first pass but checks the suitability of the following individuals and on the second run of the parade make a firm confident identification.

Furthermore, the 'hybrid model' theory proposed by the author may also explain the high levels of mis-identifications elicited during the target absent parades. It may be suggested that the eye-witness uses the first showing of the entire parade to 'rank' the individuals in order of similarity to their memory of the suspect and, during the second showing of the parade, using a *relative* judgment style, identify the 'best match' parade member.

However, in general the difference in the rate of false positive identifications between the two parade styles was found to be non significant for target absent parades. Thus the eye-witnesses who viewed the V.I.P.E.R. parade appear to be equally susceptible to the additional extraneous factors which promote the identification of an innocent foil during target absent parades. These factors include the implicit priming associated with attendance at all identity parades: 'Why would they prepare an identity parade for me to view if the suspect wasn't there?' (Wells, 1993), the difficulty linked with unfamiliar face recognition (Megreya & Burton, 2008; Kemp *et al.*, 1997; Valentine *et al.*, 2007; Megreya & Burton, 2006; Newell, Chiroro & Valentine, 1999) and verbal overshadowing effects (Meissner & Brigham, 2001; Schooler & Engstler-Schooler, 1990).

The presence of the significant effect of recall from the target absent parades was interesting. As discussed earlier, there was no evidence of the sequential superiority of the V.I.P.E.R. format over the traditional format. However, there was a significant effect of recall delay, which was unexpected as it is a significant *increase* in the rate of correct rejections from just 25% when 2 hour delay compared with 60% at the 24 hour delay. This is atypical

of the published data which states a well established link between increasing recall delay and decreasing correct identifications/rejections (Deffenbacher *et al.,* 2008; Memon *et al.*, 2003; Ebbesen & Rienick, 1998).

This counter-intuitive effect of recall delay was surprising and so further examination of the participants allocated to the 2 hour recall delay condition was undertaken. Upon closer examination, it was noted that there was a greater proportion of overseas students who had been randomly allocated to this particular condition compared to any of the other conditions. 11 participants were overseas students; 9 were Asian and 2 were Zambian. Moreover, several of these students (5) had only recently arrived in the United Kingdom and had not had extensive contact with Caucasian faces prior to their arrival. It may be suggested then that the poor rate of correct rejections generated by those allocated to the 2 hour recall condition may be underpinned by the other race bias. The other race bias is a well documented, robust effect characterised by a decreased ability to recognise and identify faces, and facial expressions, of people who do not belong to your own racial group (Sporer, 2001;, Meissner & Brigham, 2001; Pedzek *et al.*, 2003; Cross, Cross & Daly, 1971; Hancock & Rhodes, 2008). Clearly, this is a confounding variable to the current research and the data would benefit from replication and more stringent controls on the allocation of participants.

It is evident from the current results that there is no significant evidence of sequential superiority associated with the use of V.I.P.E.R. identity parades for target absent parades, just as there is no significant evidence to suggest that the simultaneous format is superior for target present identity parades. Such data refutes previous claims that the sequential nature of V.I.P.E.R. may compromise the rate of correct identifications (Wells, 2008); it is suggested that this effect is due to the absolute judgment style which sequential identity parades reportedly fostered in eye-witnesses. Pozzulo & Lindsay (1995) and Memon & Gabbert (2003) assert that while the witness may observe that the image of the suspect is familiar it may be tempting to reject the image in search of a better criterion match. However, it must be noted that the procedure used in Memon & Gabbert's 2003 study was completely

sequential; the identity parade was stopped as soon as the witness made a selection. Thus the findings are not a true reflection upon V.I.P.E.R. identity parades as the procedure does not adhere to PACE Code D (Home Office, 2008) which permits witnesses to view the entire identity parade twice *before* making identification.

However, the data from the current study does not just refute these claims of limiting hit rate nor does it corroborate the data stating V.I.P.E.R. *maintains* the hit rate observed with simultaneous parades; the current findings establish a precedent for V.I.P.E.R. *superiority* for target present parades. While this data was not found to be significant in the current study it may be more firmly established with further research making use of a larger and more diverse, representative sample population.

This trend for sequential superiority did not extend, as expected, to target absent identity parades. The overall trend of the data confirms that the V.I.P.E.R. format performed worse than the simultaneous format, eliciting a greater number of mis-identifications than the simultaneous format although this difference was not significant. It is suggested that the poor performance of the eye-witnesses across the entire target absent series was facilitated by two major factors. The first factor is the implicit priming each witness associates with their attendance at an identity parade (Wells, 1993). Many witnesses appear to rationalise their attendance to view an identity as a confirmatory cue that the suspect will be present and thus inclines them towards making an identification. Secondly, as posited earlier, the provisions of PACE Code D prohibits the implementation of a *truly* sequential format whereby the witness may identify the suspect immediately upon viewing them. As PACE Code D requires each individual to be viewed twice before permitting a decision to be made the V.I.P.E.R. format appears to exist in a hybrid state between sequential and simultaneous formats and, as a result, the V.I.P.E.R. format does not completely eliminate the use of the relative judgment style. Therefore, it is suggested that both of these factors influence eye-witness performance during target absent identity parades.

It is not surprising that the eye-witnesses performed poorly over all conditions across the target absent experiments; there is a wide range of publications which report witnesses' poor ability to recognise unfamiliar faces (Megreya & Burton, 2006, 2008, Kemp *et al.*, 1997, Valentine *et al.*, 2007, Bruce *et al.*, 2001, Hancock *et al.*, 2000, Newell *et al.*, 1999). In their 1997 study Kemp *et al.* tested the recognition performance of highly experienced cashiers for matching a live individual to credit card photographs. The experiment reported a significantly high rate of false positive decisions, even when conditions were optimal for recognition. This finding was supported by Megreya & Burton (2008) who also reported no benefit to recognition when matching live target to high quality photographs.

However it is also hypothesised that there is an inherent advantage to video identity parades; several studies have reported a small but significant increase in recognition, for target present parades, and correct rejections for target absent parades when witnesses viewed both the full face and three quarter profile view of the suspect (Newell *et al.*, 1999; Hancock *et al.*, 1999; Bruce *et al.*, 1987), and is particularly true for unfamiliar faces. However, more recent investigations into the three quarter face effect have found the effect to be weak at best (Liu & Chaudhuri, 2001). Yet the data from the current study does not appear to corroborate such an effect.

The results of the current study are undoubtedly contrary to the majority of published studies in the research community. While there appears to be the beginnings of a pool of evidence suggesting sequential presentation of parade members to eye-witnesses elicits a superior hit rate for target present parades as compared to the simultaneous format (Steblay *et al.*, 2001; 2011; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Memon & Gabbert, 2003; Valentine *et al.*, 2007) the key finding is certainly the distinctly poor performance of the V.I.P.E.R. format for target absent identity parades. While there have been concerns raised over the possibility of an indirect publishing bias (McQuiston-Surrett *et al.*, 2006; Steblay *et al.*, 2001; 2011) of more pressing concern is the procedural issues on which much of the claimed evidence for sequential superiority for target absent parades is derived. As most of these studies have employed an absolute sequential procedure, as recommended by the

community, the data generated has clear validity issues. While their value is considerable as a means for supporting the adoption of a completely sequential nature their findings tell us very little about the merits of the V.I.P.E.R. procedure in its current state. It may be accurate to state that there is a superiority effect for V.I.P.E.R. during target absent parades when a *truly* sequential procedure is employed however, at this moment bound by PACE Code D provisions, V.I.P.E.R. is not a truly sequential procedure and as the current study reflects, there is no evidence that the V.I.P.E.R identity parade format significantly reduces the rate of false identifications during target absent parades.

Chapter 3:
Using ‘Special Measures’ when Testifying

Introduction

Testifying in court before the accused is possibly one of the most distressing results of witnessing crime. It has been widely accepted that having to face the defendant whilst giving evidence has far reaching effects upon witnesses, jurors and the judicial system (Lipton, 1977; Loftus, 1975; Turtle & Wells, 1988; Kebbell & Giles, 2000; Wheatcroft, Wagstaff & Kebbell, 2004; Golding, Fryman, Marsil & Yozwiak, 2003; Regan & Baker, 1998; Golding, Sego, Polley & Hasemann, 1995; Goodman, Tobey, Batterman-Faunce, Orcutt, Thomas, Shapiro & Sachsenmaier, 1998). In recent years legislative steps have been taken to improve the witness experience whilst attending court through reducing anxiety in relation to testifying. The most prominent of these advancements in Scotland was the introduction of The Vulnerable Witnesses (Scotland) Act of 2004 which allowed for a range of 'Special Measures' to enable certain witnesses to give their best quality evidence to jurors by reducing witness anxiety (The Vulnerable Witnesses (Scotland) Act, 2004; Richards, Morris & Richards, 2008).

The 'Special Measures' are available to three specific subsets of witness: all children under the age of sixteen; adults who have a mental impairment and adults who have experienced intimidation and/or severe trauma in relation to their evidence (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003, The Vulnerable Witnesses (Scotland) Act, 2004, The Youth Justice and Criminal Evidence Act, 1999). Many of the Special Measures were previously available to child witnesses but not for adult witnesses. The present study does not address the issue of child vulnerable witnesses but rather uses an adult vulnerable witness. Examination of the ethical implications associated with using child witnesses and the costs, both financial and time, were too prohibitive. Despite this, because of the relatively recent extension of Special Measures to adults there is a decided scarcity of research examining the effects of adults using such Special Measures upon jurors' perceptions, and so the current research was designed to focus solely on the adult vulnerable witness, rather than the child.

But what are these Special Measures? Many of them are already familiar to the public; use of a wooden screen, CCTV links, use of a supporter and several are perhaps less publicised; taking of evidence by a commissioner; pre-recorded video evidence and use of prior statements as evidence in chief (The Vulnerable Witnesses (Scotland) Act, 2004; The Youth Justice and Criminal Evidence Act, 1999). Of these two are most commonly employed (Richards, Morris & Richards, 2008; Criminal Law Review Editorial, 2004): use of a screen and live CCTV links. Use of a Screen is relatively simple and requires little procedural change in the trial process (Richards, Morris & Richards, 2008); before the witness is presented to the court a large wooden screen is positioned in the courtroom. This screen extends from the witness entrance, past the witness box and in front of the defendant. Thus, the witness enters, testifies and leaves the court without ever having to make eye contact with the defendant. Crucially, the screen does not impede the jury or the judge's view of the witness and, so as to not contravene the rights of the defendant, the defendant can see the witness on a television screen. The primary aim of using a screen to testify is to eliminate the necessity for the vulnerable witness to confront the defendant whilst testifying (Birch, 2000).

Live CCTV links are also commonly used in courtrooms across the country (Richards, Morris & Richards, 2008). When testifying in this manner the witness is not required to enter the court at any point. Instead they deliver their evidence from an entirely separate location; frequently another room within the courthouse but not always, via two-way CCTV cameras. The witness is shown upon the television screens within the courtroom while the witness can see judge, prosecuting and defence counsel. While there are some issues regarding the successful use of CCTV links they remain a popular choice for vulnerable and intimidated witnesses (Richards, Morris & Richards, 2008; Criminal Law Review Editorial, 2004; Hamlyn, Phelps, Turtle & Sattar, 2004).

Pre-recorded Video Evidence (PVE) is currently one of the more rarely used Special Measures, particularly by adult vulnerable witnesses (Criminal Law Review Editorial, 2004; Richards, Morris & Richards, 2008; Davies, Wilson, Mitchell & Milsom, 1995). It is most

frequently used for witnesses who experience significant anxiety in relation to testifying or who have been subject to intense intimidation, in sexual abuse cases, hate crimes and serious domestic violence (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003). These witnesses are often the *only* witness for the prosecution and often have the most potentially damaging evidence, but, frequently, the quality of their evidence is impaired through their anxiety and fears. Pre-recorded Video Evidence can completely remove the need for vulnerable and intimidated witnesses to attend the trial; if the defence counsel agrees no cross-examination is required or attends the recording to question the witness at the arranged time. It is for this reason that Pre-recorded Video Evidence is infrequently used in addition to concerns from judges regarding the fairness and effects of its introduction to court (Richards, Morris & Richards, 2008; Hamlyn *et al*, 2004).

While the introduction of the Special Measures appear to have been warmly received by witnesses and witness support services it must also be mentioned that judicial professionals continue to have concerns regarding their implementation and the effects they may exert on jurors' perceptions and trial outcomes. There has been outspoken opposition from defence lawyers as there is concern that a witness being permitted to testify using Special Measures is perceived as an implicit admission of guilt by jurors (Tausz & Ellison, 2005; Hoyano, 2001; Birch, 2000). The Home Office has commissioned a number of studies to determine the legality of the Special Measures and to examine whether there is an implicit connotation of defendant guilt attached to a witness choosing to testify using Special Measures, specifically use of a live CCTV link and Pre-recorded Video Evidence.

In examining the legality of the Special Measures permitted under current British law both Tausz & Ellison (2005) and Hoyano (2001) reviewed the Special Measures in relation to the European Convention of Human Rights (ECHR) to determine if defendants' rights are compromised. It is the opinion of both of these papers that the conditions under which the Special Measures are permitted do not contravene defendants' right to fair trial; indeed Hoyano (2001) states that use of live CCTV links may benefit the defendant by reducing the impact of witness testimony. This finding is supported by Swim, Borgida & McCoy's 1993

experimental study which reported mock jurors as being less likely to convict when testimony was delivered via CCTV link. Conversely, several experimental studies have found some evidence suggestive of a negative bias towards defendants when witnesses choose to testify using a Special Measure (Goodman *et al,* 1998; Swim *et al*, 1993). In their 1998 study Goodman, Tobey, Batterman-Faunce, Orcutt, Thomas, Shapiro & Sachsenmaier investigated the effect of CCTV testimony upon mock jurors. Their results clearly show that although the mock jurors were no more likely to convict when CCTV testimony was used there was a weak but significant negative effect upon jurors' perceptions of the defendant. Generally, as the scientific community remains divided over whether defendants' right to due process are compromised by the use of Special Measures, members of the Scottish Justiciary, particularly defence lawyers and judges (Hamlyn *et al*, 2004; Burton, Evans & Sanders, 2006; Richards, Morris & Richards, 2008), remain concerned over the issue.

Jury groups are notoriously difficult to study (McCabe & Purves, 1974; Tinsley, 2001; Pennington & Hastie, 1999); deliberations are inherently secret and free from external observation. It is because of this surety of the judicial system that research focusing on factors which may influence juror perceptions and trial verdicts relies almost entirely upon 'mock juror' studies. While some studies have attempted to use other methodologies such as the shadow jury (McCabe & Purves, 1974) or self-report measures completed by actual jurors after conclusion of the trial (Zander & Henderson, 1993; Findlay, 1994; Jackson, 1996), the use of 'mock jurors' and simulated trials is most commonly employed. Moreover, many of these studies use mock jurors in isolation; there is no deliberation phase to the experiment (Landstrom, Granhag & Hartwig, 2005; Landstrom & Granhag, 2008; Bornstein, 1999; Pozzulo, Crescini & Panton, 2008; Goodman *et al.*, 1998; Swim *et al.*, 1993, Tetterton & Warren, 2005; Golding *et al.*, 2003). There are reasons for this, primarily that it is much more economical to recruit participants on an individual basis and collect their individual perceptions of the witness. It is time-consuming and expensive to attempt to arrange a large group of participants to attend simultaneously and last minute drop-outs could lead to the experiment being scrapped due to insufficient numbers, particularly relevant in jury simulations (Bornstein, 1999; Pozzulo *et al*., 2008). Of course, there are apparent

disadvantages with this methodology; many mock juror studies have a preponderance of students acting as jurors due to their easy accessibility for researchers and, perhaps more importantly, that a key component of all jury trials is being ignored. Addressing first the over-use of students as mock jurors; this can be problematic for the resultant data. There is research which has suggested that younger adults, particularly students, are more lenient and less likely to convict than older adults or non-students (Bornstein, 1999). There are concerns that research findings derived from student populations have limited generalisability to the wider population due to the infrequency with which students serve on actual juries (Sears, 1986; Bray & Kerr, 1982). Yet a relatively simple study by Finkel and colleagues (Fulero & Finkel,, 1991; Finkel & Duff, 1991; Finkel & Handel, 1989) asked two groups of mock jurors, students and non-students, to read through a pamphlet of hypothetical trials and asked them to deliver a verdict on each trial. In each study, no difference was found between the students' and non-students verdicts'. Regarding the absence of a deliberation phase to many mock juror studies, Bornstein (1999), in his review of jury research from the past 20 years is more open on this subject. His analysis of the research has led him to conclude that while there is certainly a lack of research employing deliberation simulations, this does not mean that the published literature cannot be extended to the wider population. More research is clearly required in this area to corroborate such claims.

While there is a distinct lack of published data on the effects these Special Measures have on jurors when used by an adult witness, there is a small but established range of papers which have assessed the impact of the screen and CCTV link Special Measures when used by child witnesses. Currently, there have been no studies investigating the effect Pre-recorded Video Evidence exerts upon mock jurors perceptions of witnesses.

The data from child studies has produced suggestive evidence that when the child witness testifies via CCTV link they are more confident, consistent and provide more detailed accounts and are less anxious (Goodman *et al*, 1998; Davies & Noon, 1991; Golding *et al,* 2003; Wheatcroft *et al*, 2004; Regan & Baker, 1998). Yet, when mock jurors are asked to

view a child witness testifying in a open court with one who testifies via CCTV link, they rate the child in the open court as more credible, honest, detailed and confident than the child who testified via a live CCTV link (Ross *et al.*, 1994; Orcutt, Goodman, Tobey, Batterman-Faunce & Thomas, 2001; Swim *et* al, 1993; Goodman *et al*, 1998) however, there is also evidence that jurors are less likely to convict when the child testified via live CCTV link (Swim *et al*, 1993; Goodman *et al*, 1998) and that jurors rate testimony delivered in this manner as having less immediacy and emotional impact upon them (Davies & Noon, 1991). Yet it must be noted that there are several other factors which may influence jurors' perceptions of child witnesses. Indeed, the degree to which adult jurors trust child witnesses varies dramatically across the literature (Quas, Thompson & Clarke-Stewart, 2005; Landstrom, Granhag & Hartwig, 2007). Firstly, there is a significant percentage of the adult population who fail to recognise the limitations of children's autobiographical memory. For example, in their 2005 study, Quas *et al.* asked potential jurors and students some general questions about children's memory and testimony; e.g. "If an 8 year old child had been repeatedly sexually abused as an infant, would they remember it?". Their results were conflicting; 38% of their participants believed that children were no more susceptible to leading questions than adults were while 46% also believed that repeated use of open ended questions, such as "Tell me what happened?", led children to make more frequent false claims. This data suggested that adults both overestimate and underestimate the suggestibility of child witnesses. It is also a robust finding that adults do have a general scepticism of children's memories but are aware that younger children are more easily suggestible than older children (Quas *et al.*, 2005; McAuliff & Kovera, 2007; Laimon & Poole, 2008). Finally, adults are more likely to believe a child's testimony when they are directly involved with the alleged crime rather than a bystander (Laimon & Poole, 2008; Goodman, Golding, Helgeson, Haith & Michelli, 1987).

Despite much of the data deriving from child witnesses, the frequency with which this finding of diminished impact and lack of immediacy of testimony has been reported in the literature suggests it to be a prevalent phenomenon (Davies & Noon, 1991; Swim *et al*, 1993; Goodman *et al*, 1998; Richards, Morris & Richards, 2008; Landstrom, Granhag & Hartwig,

2005; Landstrom *et al*., 2007), to the extent that members of the judicial community are advised to caution all witnesses of the effect when considering employing Special Measures (Richards, Morris & Richards, 2008; Hamlyn *et al*, 2006). But why do jurors appear to perceive witnesses who testify from out with the courtroom as less credible and thus attach less weight to their evidence? The one school of thought on this issue is that testimony delivered from outside the courtroom limits the jurors' ability to assess the witness completely, as would be possible in an open court with no Special Measures (Richards, Morris & Richards, 2008; Hamlyn *et al*, 2006; Davies & Noon, 1991; Goodman *et al*, 1998; Landstrom *et al.*, 2005; 2008; Nisbett & Ross, 1980). There is a substantial body of evidence which clearly indicates that jurors rely not only upon the content of the witness' testimony but more importantly upon body language indicators (Boccaccini, 2002; Penrod & Cutler, 1995; Tetterton & Warren, 2005; O'Barr, 1974; Landry & Brigham, 1992; Orcutt *et al*, 2001; Landstrom *et al*, 2005), emotional demeanour (Golding *et al*, 2003) and verbal quality (Boccaccini, 2002; Wheatcroft *et al*, 2004). Certainly all of these factors are available when the witness testifies in an open court or from behind a wooden screen; however an issue arises when the witness elects to testify via CCTV link or Pre-recorded Video Evidence. Both of these presentation methods restrict the degree to which jurors may observe the witness: both formats provide only a limited view of the witness (from the shoulders upwards) thus there is no opportunity to note the complete body language of the witness. Employing a live CCTV link to testify has some advantage to jurors over Pre-recorded Video Evidence as the examination and cross is conducted live. Therefore the jurors may still observe the witness' emotional response to questioning which aids in their determination of witness credibility, accuracy and deceitfulness (Tetterton & Warren, 2005; Landstrom *et al*, 2005; Penrod & Cutler, 1995; Golding *et al*, 2003). Conversely, Pre-recorded Video Evidence removes the adversarial nature of trial entirely and thus deprives jurors of emotional response in addition to limiting body language. It appears that the cues used by jurors to assess witnesses are severely diminished when testimony is delivered from outside the courtroom leading jurors to perceive them as less credible and accurate and detracting from immediacy of the testimony and weakening the emotional impact that that witness exerts upon each member of the jury (Davies & Noon, 1991).

A related theory is that of the 'vividness effect' proposed by Nisbett & Ross in 1980. According to this theory, the more vivid the testimony, the more positively it is received and testimonies are considered vivid if they are 'emotionally interesting, concrete and imagery-provoking and proximal in a sensory, temporal or spatial way' (Nisbett & Ross, 1980, p.45). Vivid testimonies have been shown to be perceived as more credible, are more likely to be remembered and are attended to more by jurors when compared to pallid testimonies (Bell & Loftus, 1985). Presentation mode of the testimony is an obvious factor which can affect the vividness of a witness' testimony and is of particular relevance to the use of Special Measures in court. Proximal, face-to-face interaction is considered to be more vivid than a distal presentation method, such as via CCTV link or Pre-recorded Video Evidence. Under the terms of Nisbett & Ross's theory, proximal information is more emotionally involving and more intense and so it can be proposed that live testimonies are more vivid, due to their face-to-face interaction, compared to distal testimony presentation methods. Due to this increased vividness, it can be speculated that jurors are more likely to perceive live testimonies as more positive, emotional, credible and memorable.

Thus the current study aimed to address some of the persisting issues and debate surrounding the implementation of Special Measures. Of the range of Measures currently available to vulnerable and/or intimidated witnesses three were identified as being of most interest; use of a wooden screen; live CCTV link and pre-recorded video evidence. The 'use of a wooden screen' and 'live CCTV link' Measures were selected on the basis of their frequency of use; CCTV links are most commonly requested with 42.2% of the 787 identified vulnerable and/or intimidated witnesses identified by a 2008 review conducted on behalf of the Scottish Executive (Richards, Morris & Richards, 2008) while screens are employed by 22.5%. Conversely, Pre-recorded Video Evidence was selected for investigation as it is a non-standard Special Measure, used in highly stressful cases only, and is the Measure which requires the most significant change in trial procedure and early engagement by all interested parties (The Vulnerable Witnesses (Scotland) Act 2004; Richards, Morris & Richards, 2008; The Youth Justice and Criminal Evidence Act, 1999). In their 2008 review

Richards, Morris & Richards report only one occasion where Pre-recorded Video Evidence has been permitted for a single witness, an implementation rate of just 0.12%.

It is therefore apparent that although the implementation of Special Measures in Scotland, and the wider United Kingdom, has become widespread, there remains a distinct lack of scientific evidence to settle the divergent views of practitioners and psychologists in view of how using a Special Measure to testify alters jurors' perceptions of *adult* witnesses, regardless of testimony content. Furthermore there is a need for further evidence to clarify whether Special Measures compromise jurors' perception of defendant guilt and readiness to convict. To do so, the current series of experiments were designed to determine exactly how jurors perceive witnesses who choose to testify via Special Measures. Mock trials would be recorded and the mock jurors asked to rate how credible they perceive the witness to be, how accurate they feel their story to be and to deliver a verdict upon the case, guilty or not guilty.

It was hypothesised that the use of Special Measures would be associated with decreased ratings of witness credibility and believability and thus decreased conviction rates. More specifically, it was hypothesised that the ratings of the witness and conviction rates would decrease progressively as the witness became increasingly removed from the courtroom and mock jurors: generally that the ratings of the witness and conviction rates would be strongest in the Open Court condition, followed by 'Use of a Screen'; 'Live CCTV Link' and finally 'Pre-Recorded Video Evidence'. Therefore it was expected that the Special Measure 'Pre-Recorded Video Evidence' would be associated with significantly poorer ratings of witness credibility and believability and poorer conviction rates than those given to the witness who testified in the Open Court condition.

Development of Stimulus Materials

Trial Transcripts

The content of the trial testimony and cross-examination was adapted from a historical trial transcript obtained through The National Archives for Scotland. The transcript was taken from a 1998 aggravated assault trial from the High Court of Justiciary in Glasgow; all personally identifiable details were removed and substituted. In this trial, the prosecution attempted to demonstrate that the defendant committed aggravated assault when he threw a rock at the victim before physically attacking the witness. The defence contended that the defendant was not guilty of the attack; that someone else had thrown the rock and attacked the witness. The witness in this trial was not entitled to use Special Measures to testify as this trial took place before Special Measures were extended to adult witnesses. In order to mark the witness as 'vulnerable' it was developed that the witness had been subjected to intimidation in relation to their testifying. This is not an unreasonable estimation as the witness themselves testifies about the group people who were with the accused at the time of the offence and who were behaving threateningly towards them. Thus, it was determined that the witness in the trial transcript to be used would have been deemed a vulnerable adult witness and permitted the choice of testifying via Special Measures. The script used by the actors was identical for both the control condition and the experimental conditions

There is a clear documented effect that extremes of witness confidence can significantly affect jurors' perceptions of vulnerable witnesses (Golding *et al*, 2003; Cliff & Burton, unpublished manuscript) thus the transcript of the trial was manipulated so as to avoid extreme confidence or anxiety. Importantly the factual content of the testimony and cross-examination from the original transcript remained unchanged; the actors' scripts were designed to be delivered in a more naturalistic style and communicate a typical level of anxiety experienced by many witnesses attending court; anxiety about the questioning, facing the accused, giving their evidence in the best manner and being in an unfamiliar situation. However, even mild behavioural cues of nervousness are often interpreted by jurors are evidence of deceitful behaviour (Boccaccini, 2002; Aron & Rosner, 1998; Boccaccini & Brodsky, 2002) and this can be difficult to overcome.

In order to control for expressed and perceived witness confidence the parameters which define a nervous and a confident witness had to be identified. These parameters were based on the descriptors of a nervous witness and a confident witness as described by the voluminous publications of the Law and Language Society established in 1974 by William O'Barr (Conley & O'Barr, 1998; Gibbons, 1995; Conley, O'Barr, & Lind, 1978; Erickson, Lind, Johnson, & O'Barr, 1978; Lind, Erickson, Conley, & O'Barr, 1978; O'Barr, 1982; O'Barr & Conley, 1976). Witness confidence is derived through all behaviour that witnesses engage in whilst delivering their testimony and includes body language, demeanour, attractiveness and displays of status in addition to testimony content and speech patterns. The Law and Language Society have extensively examined different witness speech patterns in order to determine the effects such patterns may exert upon mock jurors. The findings from the Society are prolific and provide detailed evidence of the common verbal indicators used by jurors' to determine a witness' confidence and identify deceit and thus credibility. Firstly the results clearly demonstrated that a nervous witness is hesitant when answering questions, expresses uncertainty in their recollection of the incident and typically uses many intensifiers (very; surely), hesitation forms (uh..; err…; and well…), and hedges (sort of, kind of). In addition, nervous witnesses tend to provide brief answers which lack detail leading to frequent breaks in story formation and frequent questioning from solicitors. Finally, nervous witnesses typically tend to employ an overly polite speech style to address their questioner and try to enhance their credibility and status by using low frequency words in place of high frequency words; i.e., 'utilise' in place of 'use'.

As expected the findings of The Law and Language Society identified a confident witness as verbally the complete opposite of a nervous witness. When questioned, a highly confident witness will provide highly descriptive answers and will yield a continuous, coherent narrative interspersed with few questions. The confident witness will express confidence in the veracity of their memories of the incident, display genuine emotions and their testimony will be typified by a lack of hesitation forms, intensifiers and hedges.

However, jurors rely on much more than verbal quality and content of testimony when determining witness credibility and believability (Aron & Rosner, 1998, Boccaccini, 2002). The conscious and unconscious behaviour witnesses engage in during testimony provide the juror with valuable insight as to the internal state of the witness and allows jurors to make a more informed determination of witness credibility (Boccaccini, 2002). It has been claimed that verbal content makes up only 7% of conveyed information while tone of voice conveys 38%. The remaining 55% is conveyed through non-verbal behaviour (Mehrabian, 1981), body language, which can reinforce or contradict the verbal testimony.

This 55%-38%-7% ratio has seemingly been overstated in the literature; even Mehrabian acknowledges that this derived ratio is a blatant mis-interpretation of his research findings (http://www.kaaj.com/psych/smorder.html). Mehrabian's (1981) equation was derived from research asking participants to make judgements about the emotional state of a speaker after hearing them speak just one, single word. Mehrabian clearly states that his theory, which he did not prove in this study, applies only when the speaker is talking about their own feelings and attitudes (http://www.kaaj.com/psych/smorder.html). However, more recent research has found that non-verbal behaviour is just as important to jurors' decision making processes as verbal cues, quality and content (Aron & Rosner, 1998; Boccaccini, 2002; O'Barr, 1974; Boccaccini, Gordon & Brodsky, 2005) and the use of non-verbal behaviour as a cue to determine reliability and credibility has been shown in children as young as 2 years of age (Birch, Akmal & Frampton, 2010; Fusaro & Harris, 2008). It is very difficult to control our non verbal behaviour, therefore when a witness' verbal testimony is incongruent with their body language and tone of voice it is highly likely that a juror will recognise such disparity and may view this as evidence of deceit and unreliability. This is why witness preparation, particularly in the United States, has become a significant pre-trial activity (Boccaccini, Gordon & Brodsky, 2005; Boccaccini, 2002; Aron & Rosner, 1998), especially as intensive training can significantly improve a client's perceived credibility and honesty in the eyes of a juror (Boccaccini *et al.*, 2005; Aron & Rosner, 1998; Boccaccini, 1998).

Aron & Rosner (1998) reviewed a wide body of studies examining the effects of body language highlighted commonly observed indicators of nervousness. These findings have further been corroborated in Boccaccini's 2002 meta-analysis and as ancillary observations from the extensive Law and Language Society studies. Typically, nervous witnesses often engage in averting their eyes to minimise eye contact; a major indicator of nervousness but often perceived as deceit. In addition, nervous witness were observed to have 'fidgeted' more; constantly shifting their posture, making repeated use of self adaptor gestures (touching or moving their own body); and object adaptors such as playing with nearby items (a pencil) or tapping fingers on a desk. Again, these indicators of anxiety are often construed by jurors as signs of deceitfulness and of a conflicted internal emotional state. While all of these behavioural cues are used by jurors to make a determination of witness veracity and credibility, eye contact and gaze is commonly reported as the primary factor used to gauge credibility (Aron & Rosner, 1998; Boccaccini, 2002).

As previously seen with verbal communication, the body language of highly confident witnesses is typified as polar opposite to that of a nervous witness. The findings demonstrate that they use frequent illustrator gestures to intensify what they are saying (for example: pointing to defendant or injuries); make infrequent shifts in posture, appear relaxed and face the jury. Furthermore, highly confidant witnesses were generally found to speak in a moderately loud voice with varying pitch, lean forward slightly and, most importantly, make frequent eye contact, particularly with the jury (Aron & Rosner, 1998; Boccaccini, 2002).

However, there is some evidence that the verbal and behavioural characteristics of both nervous and confident witnesses are exhibited by a deceitful person (Porter, Doucette, Woodworth, Earle & MacNeil, 2008). In their study, Porter *et al.* (2008) examined the verbal and non-verbal behaviours exhibit by both criminal offenders and non-offender participants when repeating planned truthful and deceptive accounts of an event. Participants were videotaped whilst repeating their planned truthful and deceptive accounts and these video recordings were later analysed for behavioural patterns. Their results clearly revealed that both the offenders and non-offenders engaged in similar patterns of behaviour when they

were being deceptive. The deceptive accounts from both groups of participants had fewer details than their honest account, similar to a nervous, honest witness, and made frequent use of illustrator gestures, like a confident, honest witness. Further analysis revealed that only the offenders exhibited a significant increase in the use of self-adaptor gestures and a significant decrease in emotional behaviour when they were lying. The authors suggest that their results show evidence of offenders' awareness of the relationship between non-verbal behaviours and perceived credibility but their results plainly illustrate how the verbal and non verbal behaviours of honest witnesses, anxious or confident, can be mis-construed by jurors.

Expressed and perceived witness confidence clearly exerts a powerful influence over jurors' perceptions and decisions. Thus the original trial transcript was manipulated to avoid any overt indicators of confidence or anxiety but rather to contain some elements of witness anxiety and confidence in both verbal and non-verbal communication (Appendix 3).

The transcript required more extensive alteration to accommodate the nature of pre-recorded video evidence. This Special Measure is typically a statement of the incident given by the witness with some cross-examination by the defence. Therefore the transcript was altered to remove all of the prosecution questioning to leave a detailed, coherent account of the incident delivered in a similar manner to the full testimony.

Mock Trial Videos

Advertisements were placed around the University campus to recruit actors to portray the necessary roles within the mock trials. It was specified within the advert that all applicants should have some acting experience and be available for rehearsal and recording over the coming six weeks. After a two week recruitment period auditions were held and recorded for each of the 45 respondents and, after review, five actors were recruited for the roles of the witness (one male and one female), procurator fiscal, defence lawyer and judge. Following a briefing where the aims of the experiment and the importance of conveying the appropriate level of anxiety were explained a further two weeks were designated for rehearsal.

Throughout the rehearsal fortnight the two actors playing the role of the witness were required to attend daily sessions together while the entire ensemble would meet bi-weekly to rehearse together. The 'witness' actors were asked to attend daily for two reasons: firstly their role required much more intensive rehearsal as it included non-verbal communication in addition to learning the script and tone of delivery. Secondly, as gender differences were to be control for by having participants watch a gender specific witness it was important that both of the 'witness' actors were very closely matched in terms of script delivery and body language. There was a high degree of flexibility required to accommodate all five actors schedules but all attended at least four rehearsal sessions while the 'witness' actors attended fourteen rehearsal sessions. All the actors received payment of £6 per hour throughout rehearsal and recording.

The mock courtroom at the University of Glasgow was booked for an entire day when all the actors could attend. However, on the day of recording the 'judge' actor did not attend and thus the recording procedure was altered that only the four remaining actors would be on screen while the experimenter read the lines of the 'judge'. The actors undertook several practice sessions whilst the experimenter prepared the equipment for recording.

The mock courtroom is a smaller scale model of a real courtroom based in the Law Department at the University of Glasgow. There were four mock trial videos to be recorded; testifying in an open court with no Special Measures, use of a screen, live CCTV link and pre-recorded video evidence. The experimental conditions required additional equipment: a large wooden screen to be positioned alongside the witness box; and a computer monitor to place next to the witness box to display the witness when testifying via CCTV link and pre-recorded video evidence.

Within the court a Panasonic SDR-S70 video camera was positioned in the front row of the jury box directly opposite the witness box at a distance of ten feet. The camera was positioned at head height level with the witness. This was the sole camera throughout as the trial videos were designed to simulate the actual view from the jury box when in trial.

The mock trials were all held in accordance with current Scots Law however the videos did not depict an entire trial instead focussing on the evidence from the only witness in the case: the victim. Thus the mock trial videos were seven minutes in duration for the control (open court), use of a screen and live CCTV link conditions whilst the video where the witness testified via pre-recorded video statement was four minutes.

The Open Court condition video shows the witness testifying in the normal manner with no special measures. As the courtroom was smaller than anticipated the questioning of the witness by the procurator fiscal and defence lawyer came from off-screen. The Screen condition was recorded in exactly the same manner with the addition of a wooden screen next to the witness box. When recording the CCTV condition the video camera was adjusted to a tight focus on the witness' head and shoulders so that none of the courtroom could be seen. This video was later played back on the computer monitor and recorded from the original wide angle in order to simulate live CCTV link testimony from a separate location. Finally the witness was then recorded, using a tight focus again, making a statement of their testimony. This recording was played back on the computer monitor in the courtroom to simulate the delivery of Pre-recorded Video Evidence. Still images taken from the mock trial videos show the viewpoint of the jury throughout each mock trial video (See photo 1).

There is strong documented evidence of a gender bias within jury groups (Boccaccini, 2002; Aron & Rosner, 1998; Kramer, Kerr & Carroll, 1990; ForsterLee, ForsterLee, Horowitz & King, 2003; Quas, Bottoms, Haegerich & Nysse-Carris, 2006; Wayne, Riordan & Thomas, 2001); typically, female jurors are harsher in their judgements of male defendants, a robust finding that encompasses both criminal trials for sexual abuse and murder and civil trials. For example, in their 2001 study, Wayne *et al.* examined mock jurors' perceptions of a range of sexual harassment cases. Theirs was a 2 (harasser gender) x 2 (target gender) x 2 (juror gender) design using 408 mock jurors. Wayne *et al.* found that the mock jurors were more likely to convict the harasser when they were of an opposite gender to themselves. Therefore, to control for gender effects, each of the four mock trial videos were recorded

twice; once with a female witness and again with a male witness. Male jurors viewed the male witness testifying whilst the female jurors viewed the female witness' testimony.

Photo 1: Still photos taken from the four conditions 1. Control (Open Court), 2. Use of a Screen, 3. Live CCTV Link, 4. Pre-recorded Video Evidence.

After recording of the eight mock trials the actors were released but remained available for any necessary reshoots for the next two weeks. The resolution across all of the video clips was 380TVL. Minor editing was performed on the videos to remove some of the more intrusive background noise. Using QuickTime Pro, the audio tracks were separated and any intrusive background noise which impacted upon the voice of the witness was isolated and removed. The audio was then encoded at a rate of 320Kbps to optimise the delivered audio quality.

Focus Groups

It was thought prudent to test the quality of the stimuli before moving forward with such a large scale study. In order to draw valid conclusions regarding the effects of the use of Special Measures the confidence and anxiety of the witness as perceived by the jurors needed to be of a similar level across all conditions and, more importantly, needed to convey a midline level of confidence or anxiety, avoiding extremes of either emotion.

Therefore, a focus group was organised to view each of the trial videos and rate the confidence of the witness each time. Fifteen participants were recruited from the student population of the University of Glasgow. None of the participants had any prior experience of serving on a jury. Each participant received a response sheet (Appendix 4) printed with eight Likert scales, one for each witness video. The participants were required to mark each witness on this ten point scale; 1 being extremely nervous and 10 extremely confident. The participants were not permitted to discuss their impressions and were instructed to provide their immediate perceptions. After all eight videos had been rated the participants were encourage to voice their opinions of what they had seen and to identify what behaviour had les them to this opinion. The focus group comments were recorded on audiotape for later reference.

The results from the Likert scales and discussion gave a strong indication that there was a problem with the stimuli videos. A range of scores (3-7, ±1 Standard Deviation) had been pre-determined by the experimenter as conveying a moderate level of confidence/anxiety thus any averaged score from the focus group out with of this range would be indicative of emotional extremes which may taint the primary aims of the study. The average Likert scale score across all eight mock trial videos from the 15 focus group participants was 8.72 (SD=1.12) thus the witnesses in the mock trial videos were expressing too much confidence. Analysis of the audio taped discussions revealed that the participants felt that the witnesses were exhibiting strong, continuous eye contact and were very relaxed whilst testifying with little movement or fidgeting.

Therefore, on the basis of the results the 'witness' actors were called for re-recording while the dialogue from the remaining actors would be played back via audio tape. Before re-filming the actors listened to the focus group's comments on their performance and suggestions were made to rectify the perceived high confidence in their performances.
After editing, a second focus group of 15 participants were recruited to repeat the process of stimulus checking. The procedure used was identical to the previous focus group as was the assessment criteria for the responses.

However, the analysis of the Likert scores and discussions again identified an issue with the perceived confidence of the witness actors across all eight mock trial videos. The average confidence score for the witness actors across all of the eight videos was just 2.57 (SD=2.18) and discourse analysis highlighted a behavioural issue in that the actors appeared evasive, did not make consistent eye contact and used too many hesitation forms and intensifiers in their speech.

For a second time, the witness actors were brought back for re-filming and again listened to the audiotapes from both focus groups and the previous stimuli videos before proceeding with re-recording. There was no need to alter their behavioural mannerisms from the previous re-shoot all that was required was a slight increase in jury directed eye contact and gaze and to remove some of the hesitations in their speech.

A third focus group was arranged to test the suitability of the re-recorded mock trial videos. The procedure was identical to that employed in the previous two focus groups as was the assessment criteria. Results from the Likert scales suggested that the trial videos were conveying an appropriately moderate confidence level; the average confidence score was 6.31 (SD=1.76) across all eight videos. While the mean score was well within the pre-determined range of acceptable scores, allowing for the standard deviation of 1.76 pushed this out with the pre-determined range of 3-7 (4.58, 8.07). Upon closer examination, it was evident that one participant had scored the witness the maximum (10). Removal of this participant and their score yielded a mean confidence score of 5.86 with a standard

deviation of 1.21 from the remaining 14 participants. While the standard deviation still placed this mean outside the predetermined range of 3-7 (4.65, 7.07) it was only marginally so. It was thus decided that the additional costs and time required to re-record and re-test the stimulus materials due to such a small infringement of the predetermined range of scores would be excessive. Therefore the mock trial videos were approved for further use in the experimental study and no further adjustments and re-shoots were necessary.

The focus groups also threw up an unexpected point of interest. During the debriefings the participants were also asked open-ended questions regarding their perception of the witness' credibility and accuracy. Several of the participants felt that they were not able to give a judgement on the witness' accuracy as they did not have a factual knowledge of the event in question. In discussion with the wider focus group, it became apparent that the term 'accuracy' was difficult for the participants to grasp; many reported feeling that 'accuracy' related to the factual context of the event and that, as they did not have such knowledge, it was imprudent to determine the accuracy of the witness' testimony. Research has shown that jurors do not examine each piece of evidence in isolation and combine it after the trial concludes; instead, jurors tend to formulate a 'story' continuously through evidence presentation and testimony, assessing and analysing each piece of information and placing it in their 'story' for a coherent picture of the event (Pennington & Hastie, 1990; Hastie *et al.*, 1983; Tinsley, 2001). Without a factual framework upon which to base their perception of witness accuracy, jurors are likely to struggle with building their evidentiary storyline. Further discussion with the focus group participants revealed that they felt 'believability' of the witness was a better construct rather than 'accuracy'. 'Believability' of the witness allows the mock juror to rate their perceptions of the witness's plausibility; a rating of how strongly the juror believes the witness' 'story' of the event. It was concluded by the participants that 'believability' was based largely upon the content of the witness' testimony. Discussion of the 'credibility' measurement construct was undertaken with the focus group participants. Witness 'credibility' was widely recognised by the participants to be a rating of how trustworthy or deceitful the witness was perceived to be; a judgement based predominantly on non-verbal behaviour but also emotional cues.

Pilot Study

There is currently little research investigating the effect Special Measures exert upon jurors perception of witnesses, therefore it was considered prudent to undertake a small scale pilot study to determine if the Special Measures do have an effect upon jurors deliberations and decisions and the extent of any effect.

Participants

Thirty undergraduate Psychology students from the University of Glasgow were recruited to participate. In line with current UK jury duty guidelines each participant was required to be a UK citizen and 18 years old. The mean age of the participants was 20.86 years (SD=3.16 years) and twenty-one were female. None of the participants had any prior experience of serving as a juror. All participants received payment for their time in the form of course credits.

Materials

Participants viewed the mock trial videos on a 17 inch iMac G5 computer monitor with a screen resolution of 1600x900. Audio was played back over the internal speakers which were high efficiency 17 watt amplifiers. Participants indicated their perceptions of witness credibility, believability and confidence on a set of Likert scales and were also asked to provide a verdict decision upon the case (Appendix 5).

Design

The experiment was a within subject design where the thirty participants were randomly allocated into one of three groups, (*Open Court vs. Screen, Open Court vs. CCTV link, Open Court vs. Pre-recorded Video Evidence*). Each participant would view the control trial video before viewing one of the three experimental videos. The participants were run on an individual basis, in line with the typical methodology for mock juror simulation studies. The order of presentation was counter-balanced with a short distracter task to minimise carry-over effects. There were four dependent measures: witness credibility, witness believability

and witness confidence, all rated upon a ten point Likert Scale (1=not very..., 10=highly...). The fourth dependent measure was the verdict decision which was a dichotomous choice between 'Not Guilty' and 'Guilty'.

Procedure

The participant was seated before a 17 inch iMac G5 computer monitor (resolution 1600 x 900) in a quiet room. They were told that they are a juror in an assault case and they are about to hear testimony from a witness. The witness may testify in an open court or may choose to use one of three 'Special Measures': use of a screen, live CCTV link or Pre-recorded Video Evidence. They were instructed not to make any assumptions about the reasons why the witness is testifying using a Special Measure.

The participant viewed the Open Court video and then rated their perceptions of the witness on each of the three Likert scales for credibility, believability and confidence. Each participant was also required to deliver a verdict based upon what they had seen and heard. A short distracter task was then administered before the participant viewed a second mock trial video, one of the three experimental conditions (*Screen, CCTV, PVE*). Upon completion the participant rated the witness and delivered a verdict upon the case.

Results

The data was analysed using one way ANOVA's to determine any significant differences in witness credibility and believability ratings and confidence scores. As the verdict data was categorical, logistic regression analysis was use to further analyse the data.

Open Court vs. Use of a Screen

The mean credibility score of the witness when they testified in an Open Court was 7.00 (SD=1.155) compared with 5.40 (SD=0.843) when they used a Screen. The mean believability score was 5.70 (SD=0.949) in the Open Court condition compared with 5.40 (SD=0.843) in the Screen condition and the mean confidence rating was 5.40 (0.843) in the Open Court condition compared with 4.60 (SD=1.350) in the Screen condition (See Table 1).

The number of guilty verdicts was identical in both conditions with 80% of the participants returning a guilty verdict.

	Credibility	Believability	Confidence	Verdict
Open Court (Control)	7.00 (SD=1.155)	5.70 (SD=0.949)	5.40 (SD=0.843)	80%
Use of a Screen	5.40 (SD=0.843)	5.40 (SD=0.843)	4.60 (SD=1.350)	80%

Table 1: Mean Ratings of Witness and Percentage Guilty Verdicts.

A one way ANOVA revealed the difference in credibility scores to be highly significant ($F(1,18)=12.52$, $p<0.002$) thus the witness is perceived as significantly less credible when they elect to testify from behind a screen. The difference in believability scores ($F(1,18)=0.56$, $p>0.05$) and confidence ratings ($F(1,18)=2.53$, $p>0.05$) were both non-significant. As both conditions returned an 80% conviction rate there was no effect upon verdict decisions, despite there being a significant decline in witness credibility.

Open Court vs. Live CCTV link

The jurors mean credibility scores for the Open Court witness was 7.60 (SD=1.075) compared to 6.70 (SD= 0.949) (see Table 2) when the witness testified via a live CCTV link. The mean believability scores for the Open Court witness was 5.70 (SD=1.494) compared to 6.20 (SD=1.229) for the CCTV condition witness. The Open Court witness received a mean confidence rating of 5.40 (SD=1.265) compared with a mean confidence rating of 3.60 (SD=1.265) in the CCTV condition. 80% of the participants returned a guilty verdict after viewing the Open Court condition while the rate of guilty verdicts was 70% in the CCTV condition.

	Credibility	Believability	Confidence	Verdict
Open Court (Control)	7.60 (SD=1.075)	5.70 (SD=1.494)	5.40 (SD=1.265)	80%
Live CCTV Link	6.70 (SD=0.949)	6.20 (SD=1.229)	3.60 (SD=1.265)	70%

Table 2: Mean Ratings of Witness and Percentage Guilty Verdicts.

A one way ANOVA showed the difference in credibility scores between the Open Court and CCTV conditions to be non-significant ($F(1,18)=3.94$, $p>0.05$) as was the difference in mean believability scores ($F(1,18)=0.67$, $p>0.05$). There was a highly significant difference between the mean confidence scores in the Open Court and CCTV conditions ($F(1,18)=10.13$, $p<0.005$) thus the participants perceived the witness as significantly more confident when they testified in an Open Court rather than via a CCTV link. There was no significant difference in verdict decision between the Open Court and CCTV condition ($\chi^2=0.24$, $p>0.5$).

Open Court vs. Pre-recorded Video Evidence

The mean credibility score of the witness was 7.50 (SD=1.080) when testimony was given in an Open Court compared to a mean score of 4.60 (SD=0.966) when Pre-recorded Video Evidence was used. Mean believability scores were 6.00 (SD=1.414) in the Open Court condition compared to a mean score of 4.70 (SD=0.823) in the PVE condition. The mean confidence scores awarded to the witness in the Open Court condition was 5.10 (SD=0.876) compared to a mean score of 3.50 (SD=1.080)in the CCTV condition. The rate of guilty verdicts dramatically declined from 70% in the Open Court condition to 40% in the PVE condition (see Table 3).

	Credibility	Believability	Confidence	Verdict
Open Court (Control)	7.50 (SD=1.080)	6.00 (SD=1.414)	5.10 (SD=0.876)	70%
PVE	4.60 (SD=0.966)	4.70 (SD=0.823)	3.50 (SD=1.080)	40%

Table 3: Mean Ratings of Witness and Percentage Guilty Verdicts.

One way ANOVA's revealed a highly significant difference between the Open Court and PVE conditions mean credibility scores ($F(1,18)=40.05$, $p<0.001$) and mean believability scores ($F(1,18)=6.31$, $p<0.025$) indicating that the participants perceived the witness as significantly less credible and believable when they elect to testify via pre-recorded video evidence. There was also a significant difference in the mean confidence ratings given to the witness between the two conditions ($F(1,18)=13.24$, $p<0.002$) strongly suggesting that the participants perceived the witness as significantly more anxious when they testified via pre-re-corded video evidence. Finally, logistic regression analysis revealed that there was no significant difference in the rate of guilty verdicts between the two conditions.

Conclusions

The results from the pilot study returned some interesting patterns as a result of using Special Measures to testify in court. The data strongly suggests that witnesses are perceived as significantly less credible when they choose to testify either from behind a wooden screen or via pre-recorded video evidence. While there is evidence of a similar declining trend when a witness uses a live CCTV link to testify the difference when compared to the control condition was not significant. It was surprising that the significant decline in credibility scores was not linked to a decrease in believability ratings when a screen was employed. The analysis does identify such a pattern though when pre-recorded video evidence is used in court, believability ratings decreased significantly over the open court condition. Equally unexpected was the finding that there was no decrease in the rate of guilty verdicts when there is a similar decline in credibility scores, as seen when a wooden

screen is employed as a protective measure. However such an association was observed in the pre-recorded video evidence condition where the rate of guilty verdicts was almost halved, although this was not significant.

The confidence scores were included in the pilot study responses as a manipulation check to ensure that there was not an extreme disparity in perceived confidence of the witness between the control and experimental conditions. There were significant difference returned in the control vs. live CCTV link comparison and the control vs. pre-recorded video evidence. Both ANOVA's found that there was a significant decline in perceived witness confidence between the control condition and the experimental conditions; that the witness was seen as less confident when they elected to testify via a live CCTV link or pre-recorded video evidence. However, as none of the mean confidence scores deviated outside of the pre-determined 'moderate' range (3-7 points) it was felt unnecessary to re-work the relevant stimulus mock trial videos.

The pilot study gave clear, indicative results of there being an issue with the use of Special Measures whilst testifying in court. The original pilot study was therefore expanded onto a much larger scale to fully investigate the effects of using Special Measures exerts upon jurors perceptions of vulnerable witnesses.

Experiment 1(a) Open Court vs. Use of a Screen

Introduction

By placing a wooden screen alongside the witness box, shielding the witness from the gaze of the defendant the vulnerable witness is able to present their testimony in a coherent manner to the jury. However, there are concerns amongst defence lawyers that such a concession insinuates defendant guilt *before* testimony is even begun (Richards, Morris & Richards, 2008; Birch, 2000; Hoyano, 2001; Tausz & Ellison, 2005). Thus it would not be unreasonable to expect that the jurors would perceive a witness testifying in this manner as *more* credible and believable than if the same witness testified in an open court. However,

data from the pilot study suggests an opposite trend; that using a screen during testimony leads to decreased ratings of credibility and believability. Clearly there is a need for further study into the effect that employing a screen during testimony has upon jurors perceptions of that witness and the effect that the jurors perceptions have upon the trial outcome.

Design

The experiment was a repeated measures design in which all participants viewed both the Open Court condition and the Screen condition. The dependent measures were the ratings of witness credibility and believability on a ten point Likert scale, and verdict decision (not guilty/guilty). As the stimulus videos had been carefully manipulated to control for extremes of witness confidence, the participants were also asked to rate the witness' confidence. Any participant who rated the witness outside a pre-determined range of 3-7, ±1 SD was excluded from the analysis. The presentation order of the trial video clips was counter-balanced to control for order effects; half viewed the Open Court trial first whilst the remaining half viewed the Screen trial first.

Method

Participants

Thirty three participants were recruited to take part, predominantly from the University of Glasgow student body. All participants were required to be eligible for jury duty in the United Kingdom at the time of participating; they must be aged over 18 and a British citizen. The participants ranged in age from 18 to 45 (mean age: 22.26 years, SD=6.24 years); sixteen were female. One participant had prior experience of serving on a jury for a criminal theft case. The participants were paid for their time receiving either course credits or three pounds cash.

Materials

The mock trials were played on a 17 inch iMac G5 monitor placed in a quiet room. The participants' were seated approximately 3 feet from the screen.

Questionnaire

Each participant was provided with a questionnaire (Appendix 6) to be completed after viewing each mock trial clip. The questionnaire asked the mock juror to rate how credible and believable they find the witness to be on a ten point Likert scale. In addition they were asked to deliver their verdict upon the case; 'not guilty' or 'guilty'. Although the trials were held in accordance with Scots Law, the third verdict option of 'not proven' was not included, as discussed earlier. As a manipulation check, the participants were also asked to score the witness on confidence. As the confidence expressed by the witness was manipulated to give a midline rating and, as extremes of witness confidence have been shown to influence jurors' decision making process (Golding *et* al, 2003; Wells, Ferguson and Lindsay, 1981; Boccaccini, 2002; Tetterton and Warren, 2005) any participant that rated witness confidence out with the pre-determined range (3-7, ±1 SD) was excluded from the analysis. Finally, the questionnaire asked participants to provide some simple demographics (age range and gender).

Instructions to Participants

Before viewing the mock trial video clips the participants were instructed to visualise themselves as a juror in the case they are about to hear. The participants were further instructed not to make any assumptions about the reasons for the witness using a Special Measure during their testimony. As this was a within subjects design, after completing the first questionnaire the participants were given a short, ten minute distracter task to minimise carry over effects, and before viewing the second clip, they received instructions to view the second trial video clip without comparison to the previously viewed trial clip.

Procedure

Participants were run individually and were exposed to both the control mock trial and the experimental mock trial videos; this is comparable to much of the research using the mock juror paradigm. The procedure was counter-balanced to minimise carry-over effects. Participants were seated in a quiet room before a computer monitor. Before beginning the study, the participants were instructed to imagine themselves as a juror in the case they are

about to view. They were told that the witness may or may not testify using a Special Measure. They were also instructed that the use of a Special Measure did not mean anything other than the witness had met the criteria for a vulnerable witness. No further information regarding the witness' vulnerability was revealed to the jurors, just as it would not be in a live trial. The participants were instructed to view the entire video clip before completing the provided questionnaire. When the video finished the participant completed the questionnaire rating witness credibility, believability and confidence and delivered a verdict (Not Guilty/Guilty). The participant then completed a short distraction task (10 minutes) before proceeding to view the second mock trial under the same conditions.

Upon completion of the second video the participant then completed a second, identical questionnaire. The participants were then debriefed to the true aims of the study and received payment for their participation.

Results

Manipulation Check

As a result of rating the witness's confidence level out with the preset parameters (3-7 ±1SD) three of the participants (one female, two males) were excluded and the data generated from their participation was removed from the subsequent analysis. Thus the following analysis derives from thirty participants, fifteen males and fifteen females.

Credibility Ratings

The ratings of witness confidence in the control, open court, condition was 5.63 (SD=1.84) while the ratings of credibility for the witness when testifying in the use of a screen condition was 5.89 (SD=1.81). Therefore the mock jurors perceived the witness who used a screen to shield themselves from the defendant's view as more credible than the same witness who testified in an open court (See Graph 4). However, further analysis using ANOVA found that there was no significant difference ($F(1,58) = 0.24$, $p>0.05$, $d = -0.14$) in mock jurors' ratings of witness credibility between the two methods of testifying. The ANOVA analysis was repeated including presentation order as a covariate to test for any carry-over effects

associated with within subjects' designs. The resultant analysis revealed there to be no significant effect of presentation order ($F(1,58) = 1.06$, $p>0.05$, $d = 0.32$).

Believability Ratings

The mock jurors rated the witness in the control condition a mean believability score of 5.8 (SD=1.92), exactly identical to the mean believability score of the witness who testified from behind a wooden screen (See Graph 4). Thus, there was no effect whatsoever ($F(1,58) = 0.00$, $p = 1$, $d = 0$) of using a screen to testify upon the mock jurors ratings of witness believability. Further analysis of the potential carry-over effects revealed that there was no significant main effect of presentation order on believability ratings ($F(1,58) = 2.54$, $p>0.05$, $d = 0.66$).

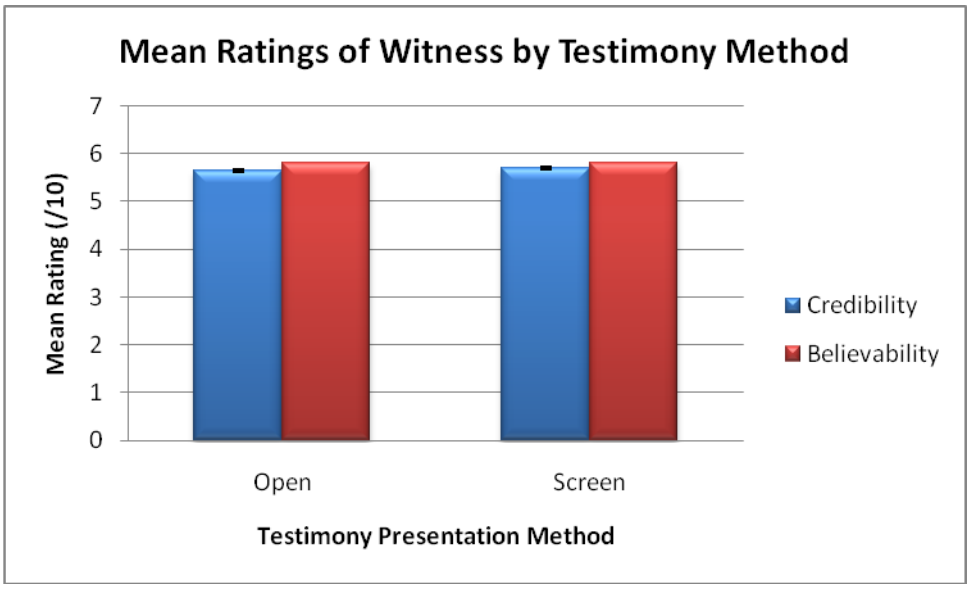
Verdict Decisions

63.33% of the mock jurors returned a guilty verdict after viewing the control mock trial where the witness testified unaided. This is comparable to the experimental, use of a Screen, condition where the conviction rate was 60%. Binary logistic regression analysis revealed that there was no significant effect of testimony presentation method upon the rate of conviction ($p>0.05$). The logistical regression analysis however did reveal there to be a significant association between the confidence scores and the likelihood of returning a guilty verdict ($\chi^2 = 60.39$, $p<0.05$, $\Phi = 0.23$) giving rise to such a regression equation:

$$\text{logit}(p) = -4.12 + 0.68 \times \text{Confidence Score}.$$

From this analysis we can say that for each 1 score increase in perceived witness confidence the probability of returning a guilty verdict increases by 19%.

The regression analysis was repeated to include presentation order as a covariate to examine any potential carry-over effects. The resultant data revealed that there was no significant effect of presentation order on the rate of guilty verdicts ($\chi^2 = 2.48$, $p>0.05$, $\Phi = 0.11$).



Graph 4: Mean Ratings of Witness by Testimony Presentation Method. Error bars show the Standard Error.

Mock Juror Gender

Although gender effects were controlled for by ensuring that mock jurors viewed a vulnerable witness of their own gender, an ANOVA was performed to determine if there were any effects associated with mock juror gender. The resultant analysis showed there to be no significant difference between the male and female mock jurors for witness credibility ratings ($F(1,58) = 2.56$, $p>0.05$, $d = 0.67$), witness believability ratings ($F(1,58) = 1.19$, $p>0.05$, $d = 0.47$), and there was no significant difference in the rate of guilty verdicts ($F(1,58) = 2.73$, $p>0.05$, $d = 0.73$).

Discussion

The data clearly indicates that there is no significant negative effect associated with using a screen to testify. Unexpectedly there was a marginal increase in witness credibility ratings when testimony was delivered from behind a screen as compared to the control condition, however this was not significant. The only negative effect found was a decrease in conviction rates in the Screen condition though this was marginal and not significant. Therefore it can be suggested from the data that a witness need not feel that the impact of

their testimony will be diminished because they have chosen to testify from behind a wooden screen rather than in an open court with no Special Measures.

Experiment 1(b) Open Court vs. Live CCTV Link

Introduction

The use of live CCTV links as a means to testify has been widely employed in British courtrooms over the last twenty years. Typically it has been most commonly associated with child witnesses in sexual abuse cases. There has been some previous research examining the effect over CCTV links upon mock jurors which has indicated a negative bias; that child witnesses are frequently seen as less credible when the testify via live CCTV link (Goodman *et al,* 1998). Yet this effect may be due to jurors' intrinsic reticence to fully trust child witnesses. As with the full range of Special Measures there is further controversy regarding the implicit connotations conveyed to jurors when a witness is permitted to use a CCTV link and to not appear live in court (Hoyano, 2001,Tausz & Ellison, 2005). As these divergent theories suggest there remains a chronic lack of data assessing the jurors' perceptions of *adult* witnesses who elect to testify via a live CCTV link.

Thus the aim of this experiment was to fully examine the relationship between jurors' perceptions of witness credibility, believability and trial verdicts and the use of a live CCTV link during testimony over testifying in an open court.

Method

Participants

Thirty one people were recruited to take part in experiment two. All thirty one participants were undergraduate students from the University of Glasgow and ranged in age from 18 years to 34 years (mean age: 21.10 years, SD=4.01 years); sixteen were female. As in the previous experiment, all participants were required to be aged 18 or over and a British citizen and received payment for their time. None of the participants had previous experience of serving as a juror.

Materials

All Materials and equipment were identical to those used in Experiment 1(a).

Procedure

The procedure employed was identical to that used in Experiment 1(a).

Results

Manipulation Check

Of the thirty one participants who completed the experiment, one female participant was excluded due to rating the witness' confidence outside of the set confidence rating parameters. The subsequent analysis was therefore based upon thirty participants, fifteen males and fifteen females.

Credibility Ratings

The mean credibility rating for the witness when testifying in the Open Court condition was 5.87 (SD=1.63) compared with a mean witness credibility rating of 5.00 (SD=1.43) when a live CCTV link was employed (See Graph 5). Therefore, the mock jurors generally perceive the witness as less credible when they testified via a live CCTV link. Further analysis revealed this to be a significant difference in credibility ratings ($F(1,58) = 4.75$, $p<0.05$, $d = 0.57$) between the two methods of testifying. The ANOVA was repeated to include presentation order as a covariate, however, the resultant analysis showed that there was no significant effect of presentation order ($F(1,58) = 2.20$, $p>0.05$, $d = 0.44$).

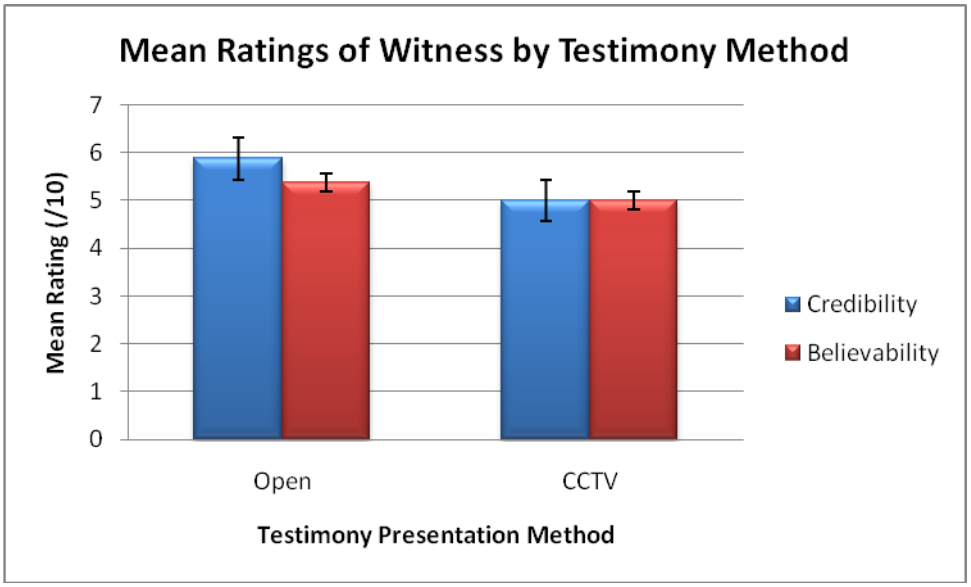
Believability Ratings

The mean believability ratings for the witness when testifying in the Open Court condition was 5.37 (SD=1.56) compared with a mean rating of 5.00 (SD=1.85) (See Graph 5). Thus the mock jurors are rating the witness as less credible when they choose to testify via a live CCTV link. Further analysis using ANOVA revealed this difference to be non significant ($F(1,58) = 0.68$, $p<0.5$, $d = 0.22$), therefore there is no significant effect of testimony method

upon mock jurors ratings of witness believability. There was no significant effect of presentation order ($F(1,58) = 1.73$, $p>0.05$, $d = 0.84$).

Verdict Decisions

56.67% of the mock jurors returned a guilty verdict in the Open Court condition compared with 53.33% in the CCTV link condition. Further analysis was undertaken in the form of binary logistic regression as the data were categorical. Results indicated that there was no significant effect of testimony presentation method upon the rate of guilty verdicts ($\chi^2 = 3.2$, $p>0.05$, $\Phi = 0.03$). The regression analysis was repeated with presentation order included as a predictor, however, the resultant data shown there to be no significant effect ($F(1,58) = 3.04$, $p>0.05$, $d = 0.52$) of presentation order upon the rate of guilty verdicts.



Graph 5: Mean Ratings of Witness by Testimony Presentation Method. Error bars show the Standard Error.

Mock Juror Gender

A one way ANOVA was performed to establish it there was any main effect of mock juror gender upon any of the dependent measures (witness credibility, witness believability and verdict decision). The ensuing data showed that there was no effect of mock juror gender upon ratings of witness credibility ($F(1,58) = 2.39$, $p>0.05$, $d = 0.44$), witness believability

(F(1.58) = 3.12, p>0.05, d = 0.61) neither was there any significant effect upon the rate of guilty verdicts (F(1,58) = 0.07, p>0.05, d = 0.5).

Discussion

The results suggest that although there is a significant effect of testimony presentation method upon mock jurors' perceptions of witness credibility the effects do not extend to ratings of believability or to the rate of conviction. Specifically when the witness chose to testify via a live CCTV link the mock jurors' rated them as significantly less credible than the witness who testified in an open court. This effect was anticipated as similar effects have been recorded for child witnesses (Goodman *et al* 1998; Swim *et al*, 1993; Davies & Noon, 1991, Landstrom *et al.*, 2005; 2007; 2008) however it was expected that the believability ratings and conviction rates would be affected in line with the ratings of credibility. The data does show some evidence of an associated decline in believability ratings and conviction rates but the differences were not significant. The significant decline in credibility ratings may be an effect of the witness becoming removed from the courtroom. It has been widely suggested (Swim *et al,* 1993; Davies & Noon, 1991; Landry & Brigham, 1992; Landstrom *et al*, 2005; 2007; Hamlyn *et al*, 2006; Nisbett & Ross, 1980) that testimony has a more significant impact upon jurors when the witness appears before them in court and thus testifying via CCTV diminishes the immediacy of the evidence. It is postulated that this effect occurs because the jurors cannot fully observe the body language and emotionality of the witness via a television screen (Davies & Noon, 1991; Nisbett & Ross, 1980; Swim *et al*, 1993), information jurors rely heavily upon when forming their opinions of the witness and deciding upon a verdict (Goodman *et al*, 1998; Boccaccini, 2002; O'Barr, 1974). However witnesses may be reassured that there appears to be no significant disadvantage associated with choosing to testify via a live CCTV link.

Experiment 1(c) Open Court vs. Pre-recorded Video Evidence

Introduction

The use of pre-recorded video evidence is certainly one of the lesser known Special Measures. Currently this Special Measure has not been widely implemented; there has been just one incident of use in Scotland since being passed into law in 2004; yet it is a measure reserved for those witnesses who often have the most potentially damaging testimony but who have a significant history of intimidation in relation to their testifying. Although there are some concerns from judges regarding its use it appears that the primary obstacle to wider implementation may be its requirement for early action (Richards, Morris & Richards, 2008). Pre-recorded video evidence necessitates preparation and organisation of prosecution and defence for cross-examination at an early stage when it is not yet certain the case will make it to trial. Due to the relatively new inception of this Special Measure, and its small uptake, it is not surprising that there has been no research examining the effect of pre-recorded video evidence upon jurors' perceptions of the vulnerable adult witnesses who use them. It is because it is reserved for the most severely intimidated vulnerable witnesses that Pre-recorded Video Evidence was included in the current trial. There are concerns about the effects it may exert upon jurors perceptions, leading them to be more sympathetic to the vulnerable witness; which may be a contributory cause to its minimal usage since the implementation of the Special Measures (Tausz & Elliston, 2005; Hoyano, 2001; Birch, 2000). Experiment 1(c) aimed to examine this Special Measure in an effort to resolve these issues and determine if the use of Pre-recorded Video Evidence does influence jurors' perceptions of witnesses and influences the trial outcome.

Method

Participants

Thirty five participants were recruited from the student body at the University of Glasgow. The age range of the participants varied from 18 years to 51 years (mean age: 28.49 years, SD=10.64 years); eighteen were female. As with the previous two experiments all participants were required to be aged 18 or over and be a British citizen in order to

participate and none of them had ever served on a jury previously. All participants were paid either £3 or course credits for their time.

Materials

All Materials and equipment were identical to those used in the previous two experiments.

Procedure

The procedure employed was identical to that used in the previous two experiments.

Results

Manipulation Check

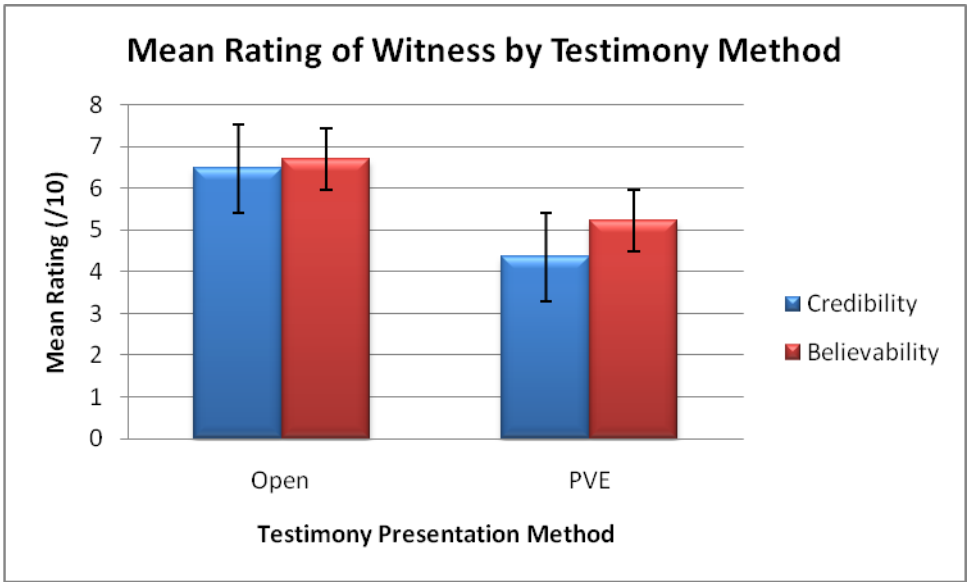
All participants rated the witness's confidence in the two clips within the range of acceptable scores. However, one female participant later withdrew from the experiment and therefore her data was excluded from the analysis below. Data analysis is based upon thirty four participants, seventeen males and seventeen females. Table 4 provides an overview of all the data.

Credibility Ratings

The mean credibility ratings when the witness testified unaided in the open court was 6.47 (SD=1.39) compared with a mean credibility rating of 4.35 (SD=1.68) in the Pre-recorded Video Evidence condition. Thus the mock jurors generally perceive the witness as less credible when they deliver their testimony in the form of a pre-recorded video statement (See Graph 6). Further analysis revealed this difference in credibility ratings to be highly significant ($F(1,66) = 31.80$, $p<0.001$, $d = 1.37$) thus there is a significant effect of testimony delivery method upon mock jurors ratings of witness credibility. The ANOVA was repeated, including presentation order as a covariate to examine any carryover effects from the repeated measures design. The resultant data showed some evidence of a significant effect of presentation order upon credibility ratings of the witness ($F(1,66) = 33.67$, $p<0.05$, $d = 0.44$).

Believability Ratings

The mean believability rating of the witness when testifying in the control condition was 6.70 (SD=1.34) compared with a mean believability rating of 5.24 (SD=1.46) when the witness testified via Pre-recorded Video Evidence (see Graph 6). Further analysis from a one way ANOVA clearly shows that this was a highly significant difference ($F(1,66) = 18.78$, $p<0.001$, $d = 1.04$) thus there is a significant effect of testimony delivery method on mock jurors ratings of witness believability. The analysis was repeated to include presentation order as a covariate and the ensuing data showed there to be no significant effect of presentation order upon believability ratings of the witness ($F(1, 66) = 2.18$, $p>0.05$, $d = 0.70$).

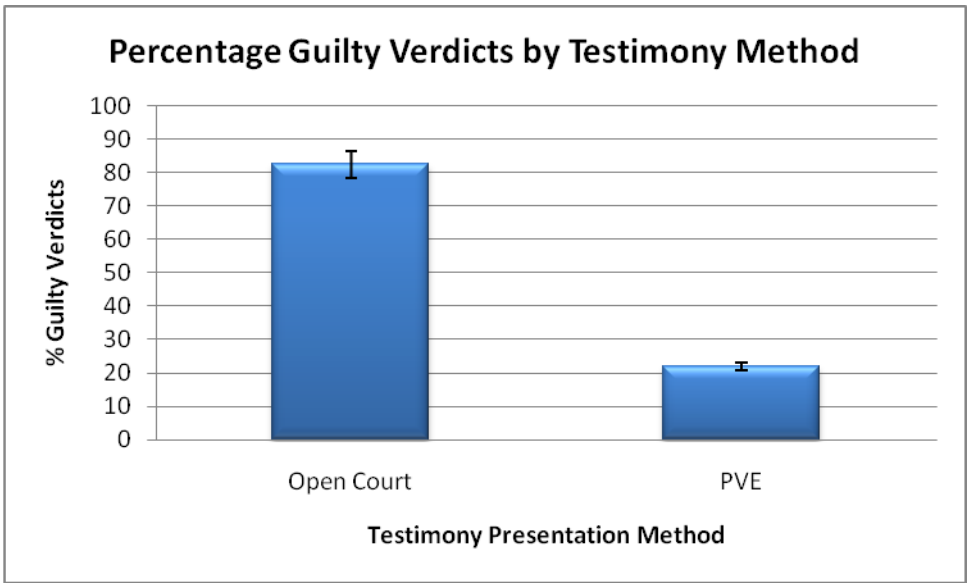


Graph 6: Mean Ratings of Witness by Testimony Presentation Method. Error Bars show the Standard Error.

Verdict Decisions

82.35% of mock jurors returned a guilty verdict after viewing the witness testify in an Open Court with no special measures. However, this rate of convictions fell dramatically to 20.59% when mock jurors viewed the experimental, Pre-recorded Video Evidence, condition (see Graph 7). As the data were categorical, further analysis in the form of logistic regression was undertaken. Results from this analysis revealed there to be a highly significant association between the testimony presentation method and the rate of guilty

verdicts (χ^2 = 61.1, p<0.001, Φ = 0.87); that those jurors who viewed the Pre-recorded Video Evidence condition were significantly less likely to return a guilty verdict compared to the jurors who viewed the Open Court condition. The regression analysis was repeated to investigate any carryover effects and so presentation order was included as an additional predictor. The data from this analysis demonstrated that there was a significant effect of presentation order upon the rate of guilty verdicts (χ^2 = 14.47, p<0.05, Φ = 0.16).



Graph 7: Percentage Guilty Verdicts by Testimony Presentation Method.

Mock Juror Gender

A one way ANOVA was performed to examine if there were any differences attributable to the gender of the mock jurors. The resultant analysis demonstrated that there was no significant differences between male and female mock jurors' ratings of witness credibility (F(1,66) = 1.70, p>0.05, d = 0.81), witness believability (F(1,66) = 0.38, p>0.05, d = 0.56) nor in the rate of guilty verdicts (F(1,66) = 0.58, p>0.05, d = 0.41).

Discussion

The results from experiment 1(c) are strongly indicative of a serious negative bias associated with choosing to testify via pre-recorded video evidence. Not only does this Special Measure significantly reduce perceived witness credibility and believability in the

eyes of the mock jurors the rates of convictions is drastically cut. The high conviction rate of 82.35% in the control condition was dramatically cut to just 20.59% of the mock jurors convicted in the pre-recorded video evidence condition. This holds significant consequences for the judicial community as a witness who chooses to testify in this manner are often those most at risk of intimidation and frequently have the most damaging testimony (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003). However, if such a vulnerable, intimidated witness chooses to testify via pre-recorded evidence they should be advised that this Special Measure may diminish the impact of their testimony in the eyes of the jurors.

There is a confounding factor in this analysis though in that there was evidence of significant order effects upon the mean ratings of witness confidence and rate of guilty verdicts. This somewhat limits the generalisability of the findings but the size of these presentation order effects were small in both instances (0.44 and 0.16) and so further research is required to clarify this issue.

Experiment Two: Effects of Deliberation

Introduction

The study of jury decision making and deliberation is, by its private nature, very difficult to achieve (Bornstein, 1999; Hastie, Penrod & Pennington, 1983; Tinsley, 2001; Davis, Kerr, Stasser, Meek & Holt, 1977; McCoy, Nunez & Dammeyer, 1999). It is impossible to gain access to actual jury deliberations in order to test hypotheses and theories without compromising the objective of the process. However, there are real concerns within the scientific community (Bornstein, 1999) that the extensive research focused upon jury manipulations, deliberations and decision making process lacks external validity as a result of this inability to observe and assess real deliberating jury groups.

While the results from the first series of experiments were indicative of a trend there is clearly an issue regarding the external validity of the data as it was generated from mock

jurors' individual perceptions and decisions. Thus in order to improve the validity of the results a second study was designed to assess both the effect the process of deliberation has on individual jurors' perceptions of a witness and whether the earlier observed effect of Special Measures occurs in a group of jurors. It was hypothesised that the introduction of the deliberation stage would have a moderating effect upon mock jurors' perceptions of the witness and the rate of convictions. Furthermore it was hypothesised, as with the earlier experiments, that the mean ratings of witness credibility and believability and rates of conviction would be poorer the further removed the witness is from the jury group.

Method

Participants

52 new participants were recruited from the student population of the University of Glasgow. The mean age was 25.46 years (SD= 4.37 years, range: 18-34 years) and 33 participants were female. The participants were randomly allocated to serve on one 13 person jury, each jury would view one of the previous four mock trial videos; open court, use of a screen, live CCTV link or pre-recorded video evidence. None of the participants had served on a jury previously. All participants were paid for their time receiving either course credit or £6.

Stimuli

The mock trial videos were the trial videos as used in the previous experiments. There were no alterations made to the stimuli. As it would have been impractical to control for gender effects thus the mock trial videos used all depicted a female witness testifying.

Materials

The participants viewed the mock trial videos on a 25 inch television screen in a large room. Seating was laid out in the fashion of the jury box; two rows of seating, one behind the other, facing the television approximately ten feet across the room. A second seating area was located in a neighbouring room which would serve as the deliberation room.

Questionnaire

All participants were given two identical questionnaires. The first asked each participant to rated their initial perceptions of witness credibility, believability and confidence and to deliver a verdict. Following completion and collection of this questionnaire the 13 member jury moved into the deliberation room. The second questionnaire was identical to the first and again asked participants to rate their perceptions of the witness after their group deliberations.

Procedure

The participants were randomly assigned to either the control, open court condition or one of the three experimental conditions; use of a screen, live CCTV link or pre-recorded video evidence. Each group of 13 participants were seated in the 'courtroom' to view the mock trial video which was displayed on a 25 inch high definition television positioned ten feet from the front row of the juror seating. They were given instructions regarding the completion of their questionnaires and the format of the experiment. Each participant then read the 'scenario setting' as described in experiment 1 before settling to view the video. Immediately after the trial video ended the participants completed the first questionnaire, no communication with the other participants was permitted. When all 13 participants had completed the questionnaire, the paper was placed into an envelope, sealed and collected by the experimenter. The participants then moved into the deliberation room where they received instructions (Appendix 8) regarding nomination of a foreperson and guidelines for verdict decisions. The experimenter was not present in the room during deliberations, however audio recording was carried out and the experimenter was seated just outside. It was explicitly stated that the group must deliberate and try to come to a unanimous verdict however, as the participants had volunteered for a 60 minute experiment, if the jury group struggled to reach a unanimous verdict the experimenter stepped in after 40 minutes to instruct that a 10-3 majority would be acceptable. When the jury had reached their verdict the foreperson notified the experimenter who returned to the deliberating room. No further conversation was permitted between participants whilst they completed the second questionnaire. The second questionnaire was identical to the first and asked the participants

to report their personal feelings about the witness and verdict after deliberation. Upon completion of the second questionnaire the participants were debriefed and received payment for their time.

Results

The results are divided into two major effects: firstly the effect of the deliberation stage on witness ratings of witness credibility and believability and rate of guilty verdicts. This first data set was analysed using a 2 x 4 repeated measures ANOVA to fully examine all the main effects and higher order effects upon witness credibility ratings and believability ratings. The data regarding the rate of guilty verdicts was analysed using multi-level logistic regression as the dependent variable was categorical. The second stage results examines the differences in witness ratings and guilty verdicts, *after* deliberation, between the control, open court condition and each of the three experimental conditions; use of a Screen, live CCTV link and pre-recorded video evidence. This data was analysed using hierarchical log-linear regression to test for all associations between the variables.

Effects of Deliberation

Credibility Ratings

The mean credibility ratings of the witness in the control condition before deliberations was 3.54 (SD=1.51) compared with a mean rating of 3.31 (SD=1.44) after deliberation. When the witness testified from behind a screen the mean credibility rating before deliberation was 4.69 (SD=1.55) compared with a mean rating of 4.00 (SD=1.29) after deliberation. The mean credibility rating before deliberation when the witness testified via CCTV link was 4.62 (SD=1.66) compared with 3.92 (SD=1.55) after deliberation. When testimony was delivered in the form of pre-recorded statements the witness' mean credibility ratings were 5.62 (SD=1.39) before deliberations compared with 5.00 (SD=1.16) after deliberations (see graph 8).

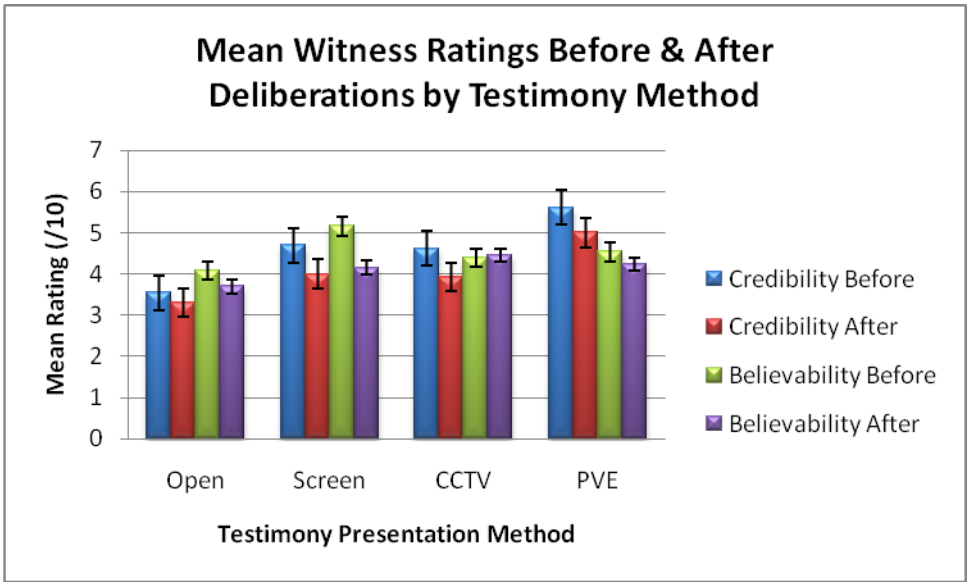
A 2 x 4 repeated measures ANOVA (Deliberation [before & after] x testimony method [open, Screen, CCTV & pre-recorded video evidence]) was performed and revealed that there was

no significant interaction term (F(3,96)=0.15, p>0.05, d = 0.43). There was no significant effect of deliberation (F(1,96)=3.85, p>0.05, d = 0.37) although this result was close to the level of significance (p=0.053). There was a highly significant effect of testimony presentation method (F(3,96)=7.35, p<0.001, d = 0.59). Planned comparisons were performed to determine which of the three experimental conditions difference significantly from the control (open court) condition. Post-hoc pair-wise comparisons revealed that the mean credibility score from the pre-recorded video evidence condition (M=1.88, 95% CI [0.80, 2.97], p<0.001, d = 0.38) was significantly different from the control condition. Neither the Screen (M=0.92, 95% CI [-0.16, 2.01]. p>0.05, 0.41) condition or the CCTV link condition (M=0.85, 95% CI [-0.24, 1.93], p>0.05, d = 0.17) differed significantly from the control, open court, condition.

Believability Ratings

The mean believability ratings for the witness in the control condition was 4.08 (SD=1.89) before deliberation compared with a mean rating of 3.69 (SD=1.70) after deliberation. Mean ratings before deliberation for the Screen condition was 5.15 (SD=2.19) compared with a mean rating of 4.15 (SD=1.68). The mean believability ratings before deliberation in the CCTV condition was 4.39 (SD=1.85) compared with an after deliberation mean rating of 4.46 (SD=1.98). Finally in the pre-recorded video evidence condition the mean believability rating before deliberation was 4.54 (SD=1.61) compared with a mean rating of 4.23 (SD=1.09; see graph 8).

A two way analysis of variance was performed to determine if there was any interaction or significant main effects of deliberation or testimony presentation method. The interaction term was reported as non significant (F(3,96)=0.41, p>0.05, d = 0.24) as was each of the main effects of deliberation (F(1,96)=1.34, p>0.05, d = 0.31) and testimony presentation method (F(3,96)=0.86, p>0.05, d = 0.19). Thus we can conclude that there is no significant effect of deliberation upon mean believability scores and that there is no significant difference in mean believability ratings between the control condition and the three experimental conditions.

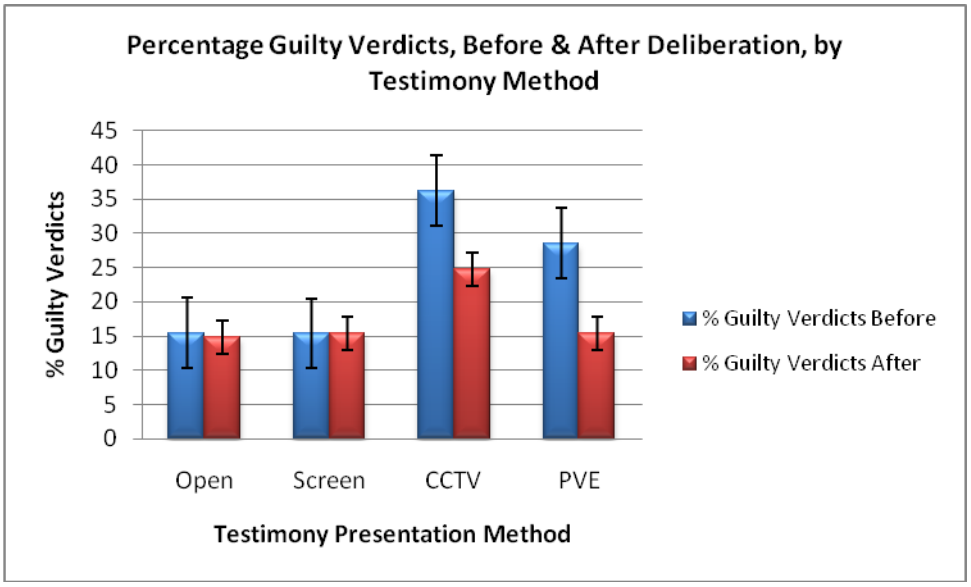


Graph 8: Mean Ratings of Witness by Testimony Presentation Method before and after Deliberation. Error bars show Standard Error.

Verdict Decisions

The rate of guilty verdicts from the Open Court condition was 15.4% before deliberations occurred compared with 14.8% after deliberations. The Screen condition yielded a conviction rate of 15.38% before deliberations which was precisely identical to the conviction rate after deliberations. 36.2% of the mock jurors who viewed the CCTV condition returned a guilty verdict before deliberations which decreased to 24.8% after deliberations. Finally 28.5% of the mock jurors returned a guilty verdict before deliberations in the Pre-recorded Video Evidence condition which decreased to 15.4% of mock jurors after deliberations (see graph 9).

Multi level logistic regression analysis was performed and revealed no significant interaction (χ^2 = 0.37, p>0.05, Φ = 0.56) and no significant main effect of either deliberations (χ^2 = 1.98, p>0.05, Φ = 0.39) or testimony presentation method ($\chi^{2=}$ 1.20, p>0.05, Φ = 0.41). Therefore while there is no evidence to suggest that there is a significant effect of deliberations upon mock jurors' verdict decisions upon the trial. Equally, there appears to be no significant effect of testimony presentation method upon the rate of guilty verdicts.



Graph 9: Percentage Guilty Verdicts Before & After Deliberation by Testimony Presentation Method. Error bars show the Standard Error.

Mock Juror Gender

Gender was not controlled for in Experiment 2 as the jury groups were comprised of both male and female mock jurors who viewed a female vulnerable witness. In order to examine if there were gender effects in the ratings of the witness and rate of guilty verdicts, a one way ANOVA was performed on the before deliberation data. The resultant analysis revealed that there was no significant main effect of mock juror gender upon witness credibility ($F(3,96) = 0.003$, $p>0.05$, $d = 0.41$), witness believability ($F(3,96) = 2.37$, $p>0.05$, $d = 0.49$) or rate of guilty verdicts ($F(3,96) = 1.59$, $p>0.05$, $d = 0.6$) *before* deliberations took place.

A second one way ANOVA was performed using the after deliberations data, however, there was still no significant evidence of an effect of mock juror gender on witness credibility ratings ($F(3, 96) = 0.96$, $p>0.05$, $d = 0.22$), witness believability ($F(3,96) = 0.68$, $p>0.05$, $d = 0.57$) or rate of guilty verdicts ($F(1,96) = 2.25$, $p>0.05$, $d = 0.64$).

Discussion

The data clearly indicates that the deliberation stage has only a weak regulatory effect upon mock jurors' perceptions and conviction rates. While the results do suggest that jurors' individual perceptions can be altered by the group dynamic, even when those initial perceptions may be relatively strong, in the current experiment there was no significant effect of deliberation upon mock jurors' ratings of the witness or conviction rate. However, there is evidence that the testimony presentation method exerts a strongly significant effect upon mock jurors' credibility ratings of the witness; when they elect to testify via pre-recorded video evidence the mock jurors' perceive them as significantly more credible as compared to the control condition.

Special Measures – After Deliberation Comparisons to Control Condition

The second stage of results examines the difference in witness ratings and conviction rates between the control condition, testifying in an open court, and each of the experimental conditions to determine if the results suggest an implicit disadvantage associated with any of the testimony presentation methods. Therefore the subsequent analyses are based solely upon the mock jurors' *after deliberation* ratings of the witness and verdict decisions, an overview of which may be seen in Table 4.

Condition	Mean Credibility Score	Mean Believability Score	Official Verdict	% Guilty Verdicts	Length of Deliberations	Majority Verdict Splits
Open Court	3.31 (SD = 1.44)	3.69 (SD = 1.70)	Not Guilty	14.8%	37 minutes	11-2
Screen	4.00 (SD = 1.29)	4.15 (SD = 1.68)	Not Guilty	15.38%	23 minutes	Unanimous
CCTV Link	3.92 (SD = 1.55)	4.46 (SD = 1.98)	Not Guilty	24.8%	44 minutes	10-3
PVE	5.00 (SD = 1.16)	5.00 (SD = 1.09)	Not Guilty	15.4%	48 minutes	11-2

Table 4: Official verdict, deliberation length and verdict splits and after deliberation means of witness and verdict decisions.

Deliberation Data

The length of time it took each jury group to reach a verdict from entering the deliberation room was recorded. Each group's deliberations was recorded via audiotape and played back later.

The Open Court jury group (9 females, 4 males) took 37 minutes to reach a majority verdict, 11-2 not guilty. The recording of the deliberations revealed that immediately after general introductions (5 minutes), a poll was taken to observe the current feelings within the group. This poll showed that 5 of the 13 jurors felt that the defendant was guilty whilst the remaining 8 were persuaded that the defendant was not guilty. After the poll, the group then engaged in discussion about the facts of the case and the behaviour of the witness they had heard from. A large part of this discussion (29 minutes) involved the majority group questioning the minority about their views and attempting to coerce them to change their opinions. During this discussion, one of the minority jurors suggested that they felt the 'Not Proven' verdict option would be more appropriate as they felt that the defendant was guilty but that there was a lack of evidence. However, as the jurors had been instructed that 'not proven' was not a permitted verdict option, this juror felt strongly resistant to delivering a 'not guilty' verdict. At this point the experimenter entered the deliberation room and announced that a majority verdict of 10-3 would be sufficient to conclude the trial. A second poll was suggested by the foreperson and taken at 37 minutes, which revealed an 11-2 split. The foreperson notified the experimenter who then re-entered the deliberation room and administered the second individual questionnaire.

The Screen jury group (8 females, 5 males) took 23 minutes to reach a unanimous verdict; not guilty. Again, the jury foreperson organised an immediate poll of the group's verdict opinions which revealed a 12-1 split in favour of 'not guilty'. The foreperson then proceeded to ask the single juror to explain why they felt the defendant was guilty. The juror spent 14 minutes explaining their opinions; that the witness was using a screen so was obviously scared of the defendant and that the witness was emotional when challenged by the defence lawyer. The majority of the jury group then challenged these assertions and suggested that

there was not enough conclusive evidence to determine the defendant's guilt. The foreperson then called for a second poll, which was unanimously in favour of a not guilty verdict. The experimenter was informed and the second questionnaire administered.

The jury group (5 females, 8 males) who viewed the witness testimony via CCTV link required 44 minutes to return a majority verdict of10-3 in favour of a not guilty verdict. This group was also different to all others in that the foreperson began the deliberations by asking each juror to offer their thoughts on the evidence and their perceptions of the witness. This initial discussion took 20 minutes, after which the foreperson organised a vote to assess the feelings of the group. The vote reflected a 9-4 split in favour of a not guilty verdict. At this point the majority of the group began to question and challenge the views of the minority, particularly focussing on the minority's assumption that the defendant must be guilty because the witness was too frightened to even enter the courtroom to testify. The foreperson was especially strident in challenging this assumption, repeatedly stating that there may be other reasons for the witness to testify via live CCTV link. After 38 minutes of deliberation, a second poll was taken and showed that there had been no change in the split; 4 jurors still voted for a guilty verdict. At this point, the experimenter entered the room and told the jurors that a majority verdict of 10-3 would be sufficient to end deliberations. Further discussions were undertaken between the majority and minority groups with much issue being made of the criteria for which a live CCTV link may be employed by a witness. Three jurors of the minority group still appeared to be resistant to returning a not guilty verdict as they believed the defendant to be culpable. After 43 minutes of deliberations, one of the four minority jurors indicated that they were willing to return a not guilty verdict and the foreperson took a final poll of each juror's verdict. Deliberations ended with a 10-3 majority in favour of a not guilty verdict before the second questionnaire was administered by the experimenter.

The jury group (5 females, 8 males) who had viewed the Pre-recorded Video Evidence condition took 48 minutes to deliver a majority verdict of 11-2 in favour of a not guilty verdict. The foreperson began the deliberations with an immediate poll of the jurors to establish the

split before engaging in any deliberation of the evidence or witness testimony. The result of this poll was split with 11 jurors favouring a not guilty verdict. The 2 jurors in favour of the guilty verdict were very resistant to agreeing a not guilty verdict. It became apparent throughout the following discussion that these 2 jurors felt that the fact that the witness had pre-recorded their testimony to avoid attending court and coming into chance contact with the defendant was highly suspicious and concluded that there must be some cause for this choice. One of these two jurors even suggested that the witness being permitted to use such a Special Measure suggested that the judge felt that the defendant was guilty and that the witness required protection from the defendant. The jurors in the majority group attempted to challenge these views; again, one juror stated that there could be many reasons why the witness had elected to testify via pre-recorded evidence whilst other jurors discussed the behaviour of the witness and the emotions they exhibited. Five jurors in the majority group were particularly concerned with the lack of emotion displayed by the witness when recounting the event and felt that their description of the injuries sustained was too matter-of-fact and impersonal. This period of discussion was briefly interrupted when the experimenter entered to specify the majority need to secure a verdict. The foreperson then asked the 2 minority jurors if they now felt differently and were able to give a not guilty verdict. When they declared themselves unable to, a final poll was taken and a majority not guilty verdict was carried with 11 jurors in favour. The experimenter was informed and the second questionnaire administered.

Credibility Ratings

Mean credibility ratings were 3.31 (SD=1.44) for the control condition compared with 4.00 (SD=1.29) for use of a Screen; 3.92 (SD=1.55) for CCTV link and 5.00 (SD=1.16) for pre-recorded video evidence (see graph 10).

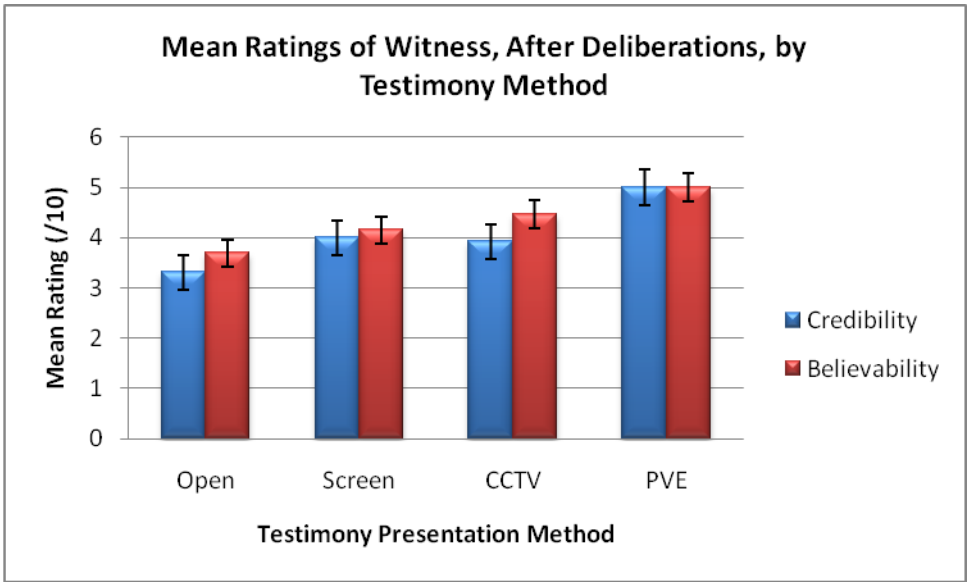
Believability Ratings

The mean believability ratings for the witness when testifying in an open court was 3.69 (SD=1.70) compared with a mean rating of 4.15 (SD=1.68) when testifying from behind a

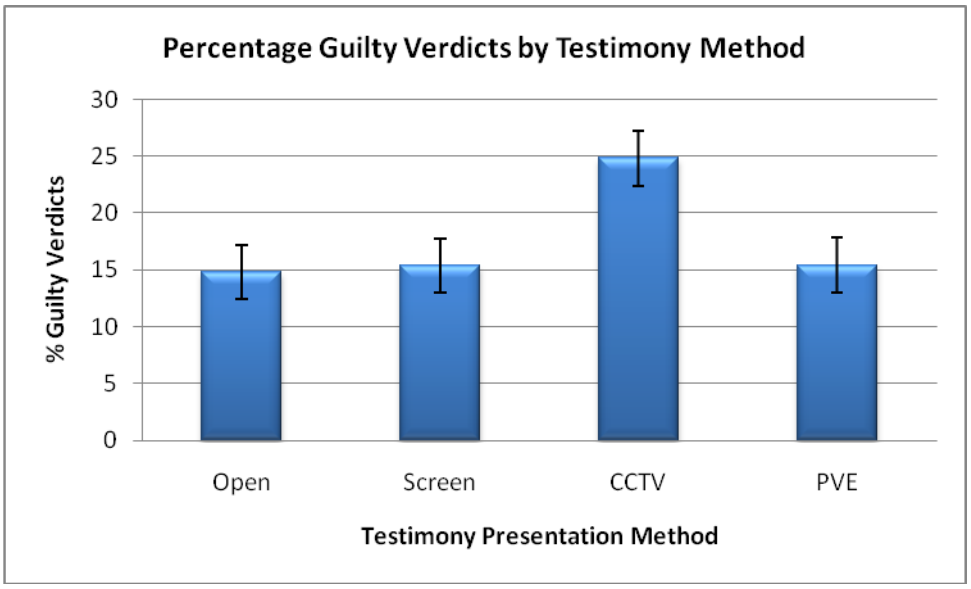
screen, 4.46 (SD=1.98) via CCTV link and 5.00 (SD=1.09) when testifying via pre-recorded video evidence (see graph 10).

Verdict Decisions

14.8% of mock jurors returned a guilty verdict in the Open Court condition compared with a conviction rate of 15.38% when the witness testified from behind a Screen, 24.8% when using a CCTV link and 15.4% when testimony was delivered via Pre-recorded Video Evidence (see graph 11).



Graph 10: Mean Ratings of Witness After Deliberations by Testimony Presentation Method. Error bars show the Standard Error.



Graph 11: Rate of Guilty Verdicts After Deliberations by Testimony Presentation Method. Error bars show the Standard Error.

Further analysis was conducted by means of hierarchical log-linear analysis. A saturated model was constructed in which each participant's *after deliberation* responses were entered as variables (*Credibility, Believability, Confidence and Verdict*) along with the experimental condition they were allocated to (*Open, Screen, CCTV, PVE*), the length of the deliberations and the type of verdict decision (*Unanimous* or *Majority*). Components of this model were then removed using backward elimination; probability for removal was $p<0.05$. The resultant analysis revealed there to be no significant effects associated with any of the variables. There was no evidence of a main effect of condition ($\chi^2 = 1.27$, $p>0.05$, $\Phi = 0.32$), credibility ratings ($\chi^2 = 0.34$, $p>0.05$, $\Phi = 0.15$), believability ratings ($\chi^2 = 1.17$, $p>0.05$, $\Phi = 0.08$), verdict decision ($\chi^2 = 0.99$, $p>0.05$, $\Phi = 0.25$), age, ($\chi^2 = 0.63$, $p>0.05$, $\Phi = 0.12$), gender ($\chi^2 = 1.49$, $p>0.05$, $\Phi = 0.2$), deliberation length ($\chi^2 = 2.04$, $p>0.05$, $\Phi = 0.37$) or verdict type ($\chi^2 = 0.70$, $p>0.05$, $\Phi = 0.09$).

Thus, it appears that the use of Special Measures by adult vulenrable witnesses is not associated with any significant effects upon mock jurors' perceptions or verdict decisions.

Discussion

The data revealed an opposite trend from that observed from the first series of experiments (1a-1c). The mock jurors did tend to rate the witness in the 'use of a screen', 'live CCTV link' and 'pre-recorded video evidence' conditions as *more* credible than the witness who testified in an Open Court, however not significantly so. Furthermore the witness in each of the experimental conditions was rated as more believable than the witness who testified in the Open Court condition although again, the differences were not found to be significant. Finally, although the officially returned verdicts were all not guilty, the rates of guilty verdicts, on an individual juror basis, were all greater in the experimental conditions than in the Open Court condition but none differed significantly. Thus, when mock jurors participate in a deliberation stage it appears that there are no significant effects, certainly no disadvantage, to the impact of a vulnerable or intimidated witness's testimony. It was interesting to note that 15.38% of the mock jurors in the Screen condition reported a guilty verdict decision in their individual second questionnaire even though they agreed a unanimous not guilty verdict in deliberation. This suggests that 2 mock jurors had not been convinced by the discussions with their fellow jurors and they appeared to have publicly complied with the majority verdict decision whilst retaining their doubts about the culpability of the defendant. It was also noted that at least one of the jurors in the Open Court condition was aware of the 'not proven' verdict option as permitted under Scots Law. As discussed earlier, this third verdict option was not included in the experimental design as research has suggested that juror understanding of this verdict option is very poor and that it is frequently used as a compromise in split jury groups (Hope, Green, Memon, Gavisk & Houston, 2007). Further research in this area though should include a study which allows jurors this third verdict option as research employing this procedure is limited (Hope *et al.*, 2007) whilst applied use of the not proven verdict is increasing every year in Scottish trials (Scottish Executive, 2006). Despite this increasing use though, the 'Not Proven' verdict option remains politically contentious and there is growing public dissatisfaction with its use (http://www.siliconglen.com/Scotland/1_8.html; Bray, 2005; Duff, 1999; Davies, 1996; Barbato, 2005; http://news.bbc.co.uk/1/hi/scotland/6500541.stm). Finally, the qualitative data from the audio recordings of the deliberation sessions appear to provide some

suggestive evidence that mock jurors implicitly associate the use of a Special Measure as an indicator of defendant guilt. This was evidenced in all experimental conditions (Screen, CCTV and PVE) to varying degrees during the deliberations. This is of concern to the judicial community and further research must be undertaken to validate these conclusions. If corroborated, changes must be implemented in the judicial system, particularly in regards to the improving the jurors understanding of the criteria used to determine witness vulnerability.

General Discussion

The data from this series of experiments has yielded some interesting trends. When the mock jurors viewed the testimony and gave a verdict individually there was suggestive evidence that the witness was rated as less credible and believable in proportion to their increasing isolation from the courtroom and jurors. Equally, there was a similar declining trend in the rate of guilty verdicts, indicating that the diminishing credibility of the witness lead the jurors to find the defendant not guilty. Yet when the study was extended to incorporate deliberating jury groups this trend was reversed, with the mock jurors rating witness credibility and believability consistently greater as they became increasingly removed from the live trial proceedings. This was accompanied by an increasing trend in conviction rates; increasing from the control condition across the Screen and CCTV link conditions before falling when pre-recorded video evidence was used. However, the rate of guilty verdicts in all three of the experimental conditions (Screen, CCTV and PVE) never differed significantly from the rate observed in the Open Court condition.

The initial run of experiments provided some strong evidence regarding the effects of using a Special Measure to testify upon mock jurors' perceptions of vulnerable and/or intimidated witnesses and their perception of defendant guilt. There is some division in the literature as to whether the use of Special Measures in court diminishes the impact of witness testimony upon jurors and thus may alter the trial outcome (Goodman *et al*, 1998; Orcutt *et al*, 2001; Swim *et al*, 1993; Davies & Noon, 1991). In addition, there are vociferous concerns from

defence lawyers and judges that the use of Special Measures implicitly suggests defendant guilt to jurors *before* they heard testimony (Richards, Morris & Richards, 2008; Tausz & Ellison, 2005; Hoyano, 2001; Birch, 2000; Hamlyn *et al*, 2006). The results from Experiments 1a-1c confirm the purported trend of decreasing testimony impact with increasing isolation of the witness from the jurors (Davies & Noon, 1991; Nisbett & Ross, 1980; Swim *et al*, 1993).

The ratings of witness credibility and believability were not significantly different in the Screen and CCTV conditions from the control, Open Court, condition. However there was a decreasing trend with witnesses using a CCTV link to testify rated lower than those who employed a screen who, in turn, were rated as less credible and believable than the witness who testified in a open court. The only significant difference in witness credibility and believability ratings between the Open Court condition and a Special Measures condition occurred when mock jurors viewed the Pre-recorded Video Evidence Special Measure. Under these conditions the witness was rated significantly less credible and believable than the witness who testified without Special Measures in an open court. Thus this finding continued the existing trend of perceived witness credibility and believability decreasing in line with declining witness contact with the jurors. In addition, further evidence of this diminishing trend is provided by the rate of guilty verdicts which exhibited a decreasing as the testimony presentation method became increasingly isolated. It should be noted that that there was no significant difference in conviction rates between the control condition and the Screen and CCTV condition however there is a decreasing trend culminating in a highly significant difference in conviction rates between the control condition and pre-recorded video evidence.

Therefore it can be concluded that witnesses who testify using Special Measures which remove them from the mock jurors appears to diminish the immediacy and impact of the testimony upon mock jurors when they work individually. Specifically, the mock jurors rate the witness as increasingly less credible and believable the more removed from proceedings the Special Measures takes the witness; thus the mock jurors rated the witness in the control

condition as most credible followed by 'use of a Screen', 'live CCTV link' and finally 'Pre-recorded Video Evidence' which is further associated with a dramatic cut in conviction rates. This finding corroborates data from other studies which have claimed a similar effect (Swim *et al*, 1993, Davies & Noon, 1991) and is clearly indicative of an implicit negative bias the Special Measures conveys to jurors. The findings from this series of studies may be interpreted as corroborating evidence for Nisbett & Ross's 1980 'vividness effect'. They claimed that jurors react most positively to testimony that is vivid, characterised by being emotionally interesting, thought provoking and, of most relevance here, proximity. Live testimony, delivered within the courtroom and face-to-face with jurors is more likely to be viewed as memorable and vivid and thus the more likely the witness is perceived to be credible and accurate.

Such a theory may go some way to accounting for the data trends observed in Experiment 2. It accounts for why there is no significant difference in witness ratings and conviction rates between the Open Court condition and the Screen condition; although the Screen shields the witness from the defendant's view, the jury, seated directly opposite the witness, remain able to clearly view the witness and their interactions with solicitors. The *vividness* of their testimony is not impeded by the screen and so their perceptions of the witness, while marginally diminished, are not significantly different to those awarded to the witness in the Open Court. Equally, the use of live CCTV links is associated with a non-significant decrease in perceived witness credibility, believability and a decrease in the likelihood of conviction. These results again may be rationalised in terms of the testimony vividness. Although the testimony is delivered live, the jurors no longer have a face-to-face, proximal, interaction with the witness and they have a limited view of their behaviour. However, the witness' testimony may still be perceived as 'vivid' as they respond to questioning, particularly if they exhibit an emotional response (Golding *et al.* 2003; Goodman *et al.*, 1998; Orcutt *et al.*, 2001; Regan & Baker, 1998).

It is when the witness elects to testify via Pre-recorded Video Evidence that the 'vividness effect' appears to be of most use in explaining the current data. Pre-recorded Video

Evidence was associated with a decrease in both witness credibility ratings and believability ratings, more so than either of the other experimental conditions, and the mean ratings observed in the Open Court condition, however, this was not a significant difference. Again, the proximity of the witness to the jurors is further removed; the witness is not present and the testimony is no longer delivered *live*. Furthermore, the emotional impact of their testimony is diminished because they are simply reading their prepared statement and answering a few, unchallenging questions (Wheatcroft *et al*., 2004). This further diminishes the vividness of their testimony in the eyes of the juror. Finally, to compound upon these factors, the jurors also have a limited view of the witness' behaviour which may impair their ability to assess honesty and credibility. All of these factors may account for the dramatic decrease in the rate of guilty verdicts when Pre-recorded Video Evidence is used as a Special Measure.

For the 22.5% of vulnerable witnesses (Richards, Morris & Richards, 2008) who elect to testify from behind a wooden screen the data does not suggest any significant advantage to be gained from testifying from an open court; they may be perceived as less credible and believable by jurors but this does not appear to influence the verdict. Equally the 42.2% of vulnerable witnesses who choose to employ a live CCTV link should certainly be informed that the jury are likely to view them as less credible if they pursue this method however, they should also be told that this bias is not likely to negatively impact upon the trial. Finally, while pre-recorded video evidence in not yet widely employed in Scotland it is often reserved for vulnerable witnesses who have potentially the most damaging evidence (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003) thus it is of great concern that jurors appear to be rating such witnesses as significantly less credible and believable. However it is the highly negative effect that the use of pre-recorded video evidence exerts upon jurors likelihood to convict which is most worrying. If choosing to testify via pre-recorded video evidence conviction rates fall to 20.59% from a high of 82.35% when testimony is delivered in an open court. Certainly there should be a responsibility to advise potential users of this Special Measure that there is a strong likelihood that their testimony will have only a weak impact upon jurors and the likelihood of a conviction is relatively poor.

From a defendant's point of view the results are advantageous; there is certainly no suggestion of an automatic implicit negative bias perceived by the jurors when viewing the witnesses who testified via one of the three Special Measures. Although several groups working with the judicial system retain concerns that the use of Special Measures indicate presumed defendant guilt rather than innocence (Tausz & Ellison, 2005, Hoyano, 2001, Birch, 2000) there is certainly no evidence from the first experimental series. Indeed the data suggests the opposite to be true; the conviction rates decline progressively as the witness elects to testify further from the live trial proceedings.

Although the first experimental series certainly provides some suggestive data in this under-researched area there are certainly some limitations which need to be addressed. If we examine the data from the three experiments there is an obvious anomaly arising with the ratings of the witness and conviction rates in Experiment 1(c) when jurors viewed the Open Court condition. The ratings of witness credibility and believability given in the Open Court condition in Experiment 1(c) were greater than the ratings awarded to the control witness in Experiments 1(a) and 1(b). Equally, the conviction rates in the Open Court condition in Experiments 1(a) and 1(b) were 63.33% and 56.67% respectively compared with a tremendously high rate of 82.35% in the Open Court condition of experiment 1(c). Thus the divergent witness ratings and conviction rates between the three control groups may be indicative of some residual carry-over effects in Experiment 1(c). The procedure was counter-balanced and a ten minute distracter task was employed, however, it does appear that the participants were affected by carry-over effects in Experiment 1(c).

All three experiments employed a within subjects design using individual participants to serve as mock jurors who assessed the witness and delivered a verdict independently. The data clearly lacks external validity as trials are not adjudicated by jurors working in isolation. Despite this limitation, the findings may perhaps be generalised to a different key component of the judicial system; the judge or sheriff. Whilst the number of solemn criminal trials are decreasing each year (Tinsley, 2001) the number of civil trials are on the increase. Many of these civil trials are purely monetary and there is distinct move towards non-jury trials in

which the judge of sheriff arbitrates the complaint. To date, there has been no published data which has examined the effects of using Special Measures upon the judge/sheriff's verdicts. Thus while the results are of interest it was important to extend the study to enhance the external validity and generate results which may be widely applied to the British criminal justice system. Certainly, the effects that the use of Special Measures exerts upon the decisions of judges and sheriff's is an important direction for future research.

Experiment 2 addressed these issues by using standard jury groups of thirteen participants and requiring them to deliberate to achieve a majority verdict. Furthermore each 'trial' was designed to appear as realistic as possible; standard procedures were employed and jurors received specific instructions regarding their expected participation.

The data from Experiment 2 immediately reported there to be no significant effect of deliberation upon juror perceptions of witness believability and defendant guilt. The individual jurors' ratings of the witness and verdict decision did not differ significantly after deliberations from their initial judgements reported immediately after testimony concluded. Thus this provides some further corroborating evidence that jurors tend to stick to their initial perceptions, particularly when this view is held by a majority of jurors (Stasser, Kerr & Bray, 1982). Why does the introduction of deliberating groups not appear to exert an effect upon the mock jurors? It is suggested that the jury group of 13 prompts the individual juror to use caution in expressing their opinions for fear of disapproval or rejection by the other jurors (Hastie *et al,* 1983; Moreland, Levine & Wingert, 1996; Blamey, McCarthy & Smith, 2000). There is documented evidence in the literature regarding group dynamics which suggests that individuals exhibit a desire for conformity, especially in groups of strangers (McKelvey & Kerr, 1988; Asch, 1951).

Despite finding that deliberations do not significantly alter jurors' perceptions, it is of much greater interest that the general pattern of data evidenced in the first experimental series was not replicated in Experiment 2. The data explicitly suggests that mock jurors rate vulnerable and/or intimidated as *more* credible and believable as their testimony is delivered

further from the courtroom. Generally, the mock jurors rated the witness who testified in an Open Court without Special Measures as least credible followed by live CCTV link, use of a screen and pre-recorded video evidence. Equally, believability scores revealed the jurors to again rate the witness who testified in an Open Court as least believable followed by those using a screen, live CCTV link and finally Pre-recorded Video Evidence. Conviction rate data revealed a marginally different pattern to that observed for the ratings of the witness; conviction rates were weakest in the control condition and increased when the witness used a screen and a live CCTV link. However, when the witness testified via Pre-recorded Video Evidence the rate of conviction declined to a similar level achieved when testimony was delivered from behind a screen; a level of conviction only slightly greater than that generated in the Open Court condition. Therefore, although the mock jurors perceive witnesses who testify via Pre-recorded Video Evidence as more credible this does not appear to affect the rate of conviction.

While this finding appears to be a positive step for witnesses there is certainly cause for concern regarding mock jurors' perceptions of *defendants* and their right to presumed innocence. Although the data reveals a general increasing trend in ratings of witness credibility and believability it must be noted that the ratings were lower overall than those given throughout experiments 1(a) – (c), when the jurors were tested in isolation, and the likelihood of a guilty verdict was also weaker than in the previous experiments. Certainly the results indicate that vulnerable and intimidated adult witnesses should not be discouraged from choosing Special Measures to aid their testimony delivery; indeed if they are eligible there is strong evidence to guide them towards the increasing isolated methods of live CCTV link and pre-recorded video evidence. Both of these conditions are associated with specific benefits for witnesses; increased credibility ratings when using pre-recorded video evidence and a greater chance of defendant conviction when using a CCTV link. For these very same reasons the data raises some concerning side effects for defendants and their legal teams. The results are strongly indicative of a negative biasing effect of the use of Special Measures upon jurors' immediate perceptions of the defendant, as widely predicted by the Justiciary and the literature (Hamlyn *et al*, 2006; Richards, Morris & Richards, 2008; Davies

& Noon, 1991; Nisbett & Ross, 1980; Landstrom *et al.*, 2007; Swim *et al*, 1993; Goodman *et al*, 1998).

The results of Experiment 2 were expected to follow a similar trend as observed in the initial experimental series; that the mock jurors would rate the vulnerable witness as increasingly less credible and believable as the Special Measure employed increasingly isolated the witness. In turn, because of these diminishing ratings the conviction rates were expected to decrease progressively. Yet the data shows an unmistakable opposite trend of *increasing* witness ratings and conviction rates.

As the only change in procedure between Experiments 1 and 2 was the change to include deliberating jury groups, it is thought that it is this change which prompted the reverse pattern. It is suggested that requiring the jury to deliberate resulted in several jurors to posit the theory that "there must be a *reason* why this witness is afraid to face the defendant" which is then pursued by the group until a majority verdict is reached. Such a line of thought obviously leads to a bias in jurors' perception of the witness and the guilt of the defendant. However, the rates of guilty verdicts overall in the second experiment were consistently low, the highest achieved being 24.8%. Evidence from the audio recordings of the deliberation phase confirmed that in each of the experimental groups, Screen, CCTV and PVE, at least one juror adopted the rationale that the use of a Special Measure indicated defendant guilt. Yet these jurors were consistently in the minority in believing the defendant guilty. A related issue in this point may be that the other jurors who adopted a 'not guilty' stance did believe the defendant to be culpable but felt that the prosecution had not met the burden of proof. It could be that, had the 'Not Proven' verdict, uniquely available under Scots Law, been an option to the mock jurors, the data may have yielded some different results. Thus it is suggested that when the mock jurors *know* they have to deliberate in a large group of strangers each individual juror becomes more conservative in their perceptions and appears to seek conformity within the jury group (Asch, 1951; McKelvey & Kerr, 1988; Hastie *et al*, 1983; Moreland, Levine & Wingert, 1996).

Despite both experiments returning diametrically opposite data the findings from Experiment 2 are strongly valid. There can be little suggestion that the study lacks ecological validity; each mock trial was recorded in a model courtroom; standard jury instructions were given; required to deliberate: first to try for unanimous verdict then a majority if necessary and standard sized jury groups were recruited, all of whom were strangers to each other. Certainly there is some degree of simulation in that the jurors did not view the trial live, just a video recording, which may raise some concerns. However, recent studies have found there to be no significant difference in results obtained from videotaped stimuli compared to live stimuli (Pozzulo, Crescini & Panton, 2008; Bornstein, 1999). In addition there may be some reservation regarding the demographics of the participants recruited to act as jurors throughout the experimental series. Certainly the vast majority of the participants in this series of experiments were university students of wide ranging age and status yet there is substantial evidence in the literature that students are consistently more lenient in their verdicts and less likely to convict (Devine, Clayton, Dunford, Seying & Pryce, 2001; Nemeth, 1981; Caspar, Benedict & Perry, 1989). Future research should address this issue by recruiting from a wider population.

However, the data from the entire experimental series has shed some light upon the effects the use of Special Measure exerts upon mock jurors perceptions of vulnerable and/or intimidated witnesses and defendant guilt. There has been little research conducted in this area especially using adult vulnerable witnesses and it appears to yield contentious results. Certainly, under the current format there is strong evidence to suggest that juries routinely perceive witnesses as *more* credible and believable as the witness become more removed from the live trial proceedings.

This is in contrast to much of the available data which claims the opposite effect; that testimony impact is diminished and witness credibility decreases with increasing isolation (Davies & Noon, 1991; Swim *et al*, 1993; Nisbett & Ross, 1980; Landstrom *et al.*, 2005, 2007). Equally, the rate of convictions increased in line with the ratings of the witness which

suggests that jurors implicitly associate use of Special Measures with defendant guilt. Clearly there is evidence to support the concerns of defence lawyers and judges.

The use of Special Measures whilst testifying remains an option only for those adult witnesses who meet stringent conditions and it is often the case that these witnesses are vitally important to the case and may have the most damaging testimony. There is clear evidence that jurors are more receptive to witness who elect to testify from out with the courtroom; they rate witnesses as most credible when using pre-recorded video evidence yet it is the use of live CCTV links which are associated with the greatest likelihood of conviction. While the introduction of Special Measures is certainly a step forward in improving the witness experience at court this advancement cannot be made at the expense of the rights of the defendant and further investigations are required to more fully understand the nature of this effect and the implications for defendants.

Chapter Four:
Summary, Policy Recommendations and Conclusions

This thesis has examined the pathway of vulnerable witnesses through the criminal justice system; from identity parade to testifying in court using Special Measures. This pathway has been radically altered over the course of the past ten years with the introduction of a video-based identity parade protocol and the implementation of the 'Special Measures' to enhance the in-court experience of the witness whilst improving testimony quality for jurors. While there is extensive literature professing the benefits of the sequential video format over the traditional simultaneous parade there is an obvious knowledge gap in that the conditions under which sequential superiority occurs remain unknown. Equally, the provisions now permitted to aid vulnerable witnesses during testimony have generally been widely welcomed (Richards, Morris & Richards, 2008; Hamlyn *et al.*, 2004; Birch, 2000; Tausz & Ellison, 2005; Hoyano, 2001), yet there has been little research into the effect their implementation has had upon jurors and no research has employed vulnerable *adult* witnesses. Therefore, it has been the primary aim of this thesis to investigate the efficacy and value of each of these key markers of a witness's journey from identification to trial.

It is apparent from the literature that there is a generally accepted theory with the research community that sequential presentation of identity parades are linked with a significantly superior rate of correct rejections during target absent parades, termed the Sequential Superiority Effect, whilst maintaining a high rate of correct identifications for target present identity parades (Steblay *et al.*, 2001; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Wells, 2008, 1993; Valentine, 2006; Pike *et al.*, 2000; Levi, 1998; Kassin *et al.*, 2001; Gronlund, 2004). However, it is also evident that much of the data supporting this effect derives from studies using an entirely sequential procedure, whereby witnesses may view each parade member once only (Steblay *et al.*, 2001; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Wells, 2008, 1993; Valentine, 2006; Pike *et al.*, 2000; Levi, 1998; Kassin *et al.*, 2001; Gronlund, 2004; McQuiston-Surrett *et al.*, 2006; Memon & Gabbert, 2003). Such a procedure, in the United Kingdom, represents a contravention of Code D of The Police and Criminal Evidence Act (1984, 2008) which stipulates that eyewitnesses must view the entire parade at least *twice* before they can make a decision (Valentine, 2006; The Police and Criminal Evidence Act 1984, Code D 2008).

The current experiments were markedly different from much of the current literature in that they were specifically designed to adhere to the existing British legislation. Therefore, all participants were instructed not to make an identification decision until the identity parade had been shown twice. This may be a contributing factor to the differing trend reported in the current analyses. Whereas much of the current literature suggests that the more traditional, simultaneous identity parade format is superior to the sequential format for target present parades (Steblay *et al.*, 2001; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Memon & Gabbert, 2003; Wells, 2008, 1993; Valentine, 2006; Pike *et al.*, 2000; Levi, 1998; Kassin *et al.*, 2001; Gronlund, 2004; Valentine *et al.*, 2007; Lindsay & Pozzulo, 1999) this was not the case in the current experiments. The data indicates that the sequential V.I.P.E.R. format was superior to the simultaneous format across all four conditions, eliciting a greater number of correct identification; however, subsequent analysis did not reveal a significant difference. Furthermore, the current study found no evidence of sequential superiority during target absent parades; the V.I.P.E.R. parade was outperformed by the traditional simultaneous parade in three of the four conditions and overall there was no significant difference in the rate of false positive identifications between the two parade types.

Addressing first the general trend of V.I.P.E.R. superiority for target present parades; evidence of V.I.P.E.R.'s superiority was not expected as the literature reports that the sequential parade, at best, maintains a similar hit rate associated with the simultaneous format (Steblay *et al.,* 2001; Lindsay & Wells, 1985; McQuiston-Surrett *et al.,* 2006; Kassin *et al.,* 2001) and, at worse, actually decreases correct identifications (Wells, 2008; Memon & Gabbert, 2003), resulting in more guilty individuals avoiding prosecution. Although in this study there was no significant difference in the rate of correct identifications between the two parade types, this so-called 'sequential superiority effect' is more frequently reported in association with target absent parades. However, as discussed earlier, the majority of the literature examining sequential identity parades has employed a wholly sequential presentation of static images rather than the partially sequential presentation of moving images. Data from the only study which has used V.I.P.E.R (Valentine, Darling & Memon, 2007) also reported no significant difference in the rate of correct identification, in line with

the findings of the current experiments. The data from the Valentine *et al.* study also provided more supporting evidence that strict sequential presentation is associated with a significant decrease in sensitivity for target present parades; the number of correct identifications is significantly reduced. In an applied setting, this could be construed as an increase in the number of guilty culprits going unidentified; this is particularly relevant in cases where there is a distinct lack of other, non-witness, evidence to support the allegation.

Turning to the target absent data from the current study; the data suggests that there is no significant difference in the rate of correct rejections between the two parade types. It was hypothesised that the V.I.P.E.R. parades would be associated with a significantly greater number of correct rejections compared to the simultaneous presentation. There was no evidence to support this hypothesis and so it appears that the V.I.P.E.R. parade is not significantly better at reducing mistaken identifications than the simultaneous parade.

It is suggested that the superiority trend for target present parades and the non-significant difference for target absent parades reported here are driven by the procedural differences between this study and the extant literature; while the large majority of research in this area has elected to use a fully sequential parade format, the current study was designed to mimic the exact procedure used by British police forces. While it is strongly recommended by researchers that the images within the identity parade be viewed by witnesses only once, UK legislation has not taken this step, requiring the entire parade be shown twice before permitting witnesses to make their decision known. It is this difference which is thought to be the key change which is driving the unanticipated results of this study.

The literature strongly favours the judgement style theory for explaining the sequential superiority effect; that a relative judgement style is used during simultaneous parades while an absolute judgement style is associated with the sequential parade format (Lindsay & Wells, 1985; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Memon & Gabbert, 2003; Levi & Lindsay, 2001; Levi, 1998; Tredoux & Chiroro, 2006; Lindsay & Pozzulo, 1999; Kneller *et al.,* 2001; Wells *et al.,* 1998; Clark & Davey, 2005). Interpreting the current results

in terms of this theory may explain why such a pattern has emerged. It is hypothesised that, due to the two-view procedure required by V.I.P.E.R., the eyewitness is able to use the relative judgement style on their first viewing before switching and engaging an absolute judgement style. In this method, the eyewitness can compare the faces both across the parade and to their memory trace of the perpetrator and identify the 'best match' faces before engaging an absolute judgement style on the second pass; focusing on the previously identified 'best matches', possibly with a higher match criterion (Meissner *et al.,* 2005; Ebbesen & Flowe, 2002; McQuiston-Surrett *et al.,* 2006), to make a final, definitive decision; reject or accept.

This hypothesised model would account for the V.I.P.E.R. hit rate being slightly greater than that associated with the simultaneous parade; the hit rates were relatively high in both parade types however it may perhaps be that the repeated presentation of the sequential V.I.P.E.R. parade improves the ability of the eyewitness to make a decision rather than choosing to reject because of uncertainty. Evidence to support this theory may be found in the literature; Memon & Gabbert (2003) examined the sequential parade and its success in target present parades. This study employed a fully sequential procedure, unlike V.I.P.E.R., and reported that the participants who viewed the sequential parade were less prone to make an identification decision and therefore rejected the target present parade more frequently than those participants who had viewed the simultaneous parade. Perhaps the legislative requirement of viewing the V.I.P.E.R. parade twice actually reverses this trend and provides a boost to eyewitness confidence and increases the likelihood of making a positive identification.

Under the terms of this proposed 'dual-judgement style' model we would then perhaps anticipate that the sequential V.I.P.E.R. parades would also elicit, in line with the literature (Steblay *et al.,* 2001; Lindsay & Wells, 1985; Levi & Lindsay, 2001; Wells, 2008, 1993; Valentine, 2006; Pike *et al.,* 2000; Levi, 1998; Kassin *et al.,* 2001; Gronlund, 2004; McQuiston-Surret *et al.,* 2006; Lindsay & Pozzulo, 1999; Valentine *et al.,* 2007), a significantly greater rate of correct rejections during target absent parades. The first pass

viewing would allow the eyewitness to compare each individual to their memory of the perpetrator and identify any potential matches before the switch to an absolute judgement, and tougher match criterion, leading to eventual rejection of the entire parade. Yet the data clearly shows that there is no significant difference between the simultaneous and V.I.P.E.R. parade types in terms of the rate of correct rejections with both formats eliciting a rate of approximately 52%.

It is thought that this poor performance of the V.I.P.E.R. format during the target absent parade is provoked by the same factors that elicited sequential superiority during target present parades. When the eyewitness makes tentative potential identifications on the first viewing of the V.I.P.E.R. parade the repeated presentation in the second viewing may, in combination with the implicit priming associated with attending a parade (Wells, 1993), lead the eyewitness to feel pressured to make an identification, particularly as the parade draws towards conclusion. These two factors of priming and pressure to decide may lead to an increase in the number of false positive identifications being made and, in turn, increases the possibility of wrongful convictions.

There are some suggestions in the current literature to support this; Lindsay, Lea & Fulford (1991) conducted a range of experiments to examine several 'violations' of the sequential identity parade procedure; namely repeated presentation. What they found was suggestive evidence that allowing eyewitnesses to view the identity parade twice increased the rate of false identifications. Lindsay *et al.* claimed that this rise in misidentifications arises from the increased pressure upon witnesses to choose; at the end of the first presentation the witness knows how many individuals are present and therefore as the second presentation of the parade draws towards its conclusion the witness feels pressured to make an identification, even though they have been explicitly told previously that the perpetrator may not be present (PACE Code D, 2008; Valentine, 2006; Levi & Lindsay, 2001; Steblay *et al.,* 2001; Lindsay & Pozzulo, 1999). It may be that this increase in selection pressure, in combination with the priming associated with attending an identity parade, leads to an inclination to make an identification of the 'best-match' individual, even though the witness

may not be absolutely certain (Steblay *et al.,* 2001; Memon & Gabbert, 2003; Lindsay & Wells, 1985; McQuiston-Surrett *et al.,* 2006). This pressure to select is why so much of the extant literature strongly advocates *singular* presentation of sequential identity parade images (Lindsay & Wells, 1985; Steblay *et al.,* 2001; McQuiston-Surrett *et al.,* 2006; Levi & Lindsay, 2001; Lindsay & Pozzulo, 1999; Valentine *et al.,* 2007; Tredoux & Chiroro, 2006; Wells *et al.,* 1998).

The legislation requiring double presentation may be leading towards a hybrid presentation-judgement model allowing eyewitnesses to use both relative and absolute judgements styles whilst also allowing them to know how many individuals make up the parade. In its current state V.I.P.E.R. is not a wholly sequential format and thus cannot be anticipated to elicit the same effects reported by the literature. Although in its current double presentation format V.I.P.E.R. performs significantly better than the more traditional simultaneous parade for target present parades what is of greater concern is its relatively poor performance for target absent parades.

V.I.P.E.R. was introduced by British police forces primarily as a cost-cutting device (Slater, 1995) yet the abundance of research proclaiming the power of the sequential parade to minimise false identifications was undoubtedly a substantial influencing factor. However, its performance at this task whilst constrained by the legal requirements of PACE Code D is significantly limited. In the current study, the rate of mis-identifications was 47% for both the V.I.P.E.R. and simultaneous parade formats; despite V.I.P.E.R. adhering to a double-blind administration (The Police and Criminal Evidence Act 1984 Code D 2008) and using a central database of images to improve parade fairness (Valentine & Heaton, 1999; Steblay *et al.,* 2001; Slater, 1995). Therefore, the current data clearly shows that there is no advantage to justice in using the hybrid V.I.P.E.R. format over the simultaneous procedure. Although these false identifications are no longer likely to proceed beyond this stage, unlike target absent simultaneous parades (Slater, 1995; Steblay *et al.,* 2001), the poor performance of the eyewitness at the parade has the propensity to damage the credibility of the witness in the eyes of the judicial system (Penrod & Cutler, 1995; Steblay *et al.,* 2001).

Therefore, on the basis of the current data, it is advised that the V.I.P.E.R. identity parade system is not associated with a significant reduction of mis-identifications as reported of the wholly sequential format advocated by the research community (McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Kassin *et al.,* 2001; Lindsay & Pozzulo, 1999; Lindsay & Wells, 1985; Levi, 1998; Levi & Lindsay, 2001). Furthermore, there is no significant benefit of using V.I.P.E.R. over the traditional format for target present parades. However, it is recommended that the V.I.P.E.R. format continued to be employed by national police forces as, despite its comparable performance to simultaneous parades, there are associated advantages to it use; the reduction in delay between incident and identity parade, the reduction in police costs and the ability to deliver identity parades in a locale which is more comfortable for the eyewitness, without the threat of meeting the accused. However policy makers must be aware that there is no inherent advantage of V.I.P.E.R. for target absent parades with the exception of reducing the number of wrongful prosecutions of an identified innocent foil (Valentine & Heaton, 1999) which would likely arise from a false identification from a traditional simultaneous parade.

It is also recommended that further investigations be carried out to examine the underlying processes of the V.I.P.E.R. parade format to gain a deeper understanding of why V.I.P.E.R. does not exert similar effects as a wholly sequential format. We speculate that the dual presentation of the parade images allows eyewitnesses to know how many individuals comprise the parade and permits them to use both relative and absolute judgements styles in a hybrid model which allows first for 'sorting' and then secondly identifying. It is recommended that policymakers further consider the necessity of PACE Code D and assess whether the removal of this legislative constraint may enhance the discriminability of V.I.P.E.R. for target absent parades. A non-PACE Code D condition was not included as part of the current study as the aims of the research was to examine V.I.P.E.R. in its applied format and its performance against the traditional, simultaneous format. There is already a significant body of research which has tested the effectiveness of the strict sequential procedure. Certainly future investigations should involve a direct test between the V.I.P.E.R. procedure and the strict sequential procedure, with particular emphasis on determining the

exact judgement styles engaged by witnesses during both types of parades. In addition, with recent research failing to replicate the Sequential Superiority Effect for target absent parades (Carlson, Gronlund & Clark, 2008; Valentine *et al.*, 2007) further research is required to determine the robustness of the claimed Sequential Superiority Effect. The controversies surrounding the Illinois Field Tests of sequential identity parades (Steblay, 2010, 2006; Wells, 2006; Mecklenburg, 2006) serve as a reminder that even basic methodological procedures may exert an effect upon the choosing rates of eyewitnesses and so future research should be stringent in its adherence to established methodologies where possible.

Specifically in terms of the V.I.P.E.R. procedure, the first step in future research should be the introduction of a non-PACE Code D variation, allowing the witness to view the entire parade once before making their decision. This procedure would eliminate the potential for a hybrid model which enables both relative and absolute judgements, and may have a significant effect upon the rate of correct, and mistaken, identifications. Further research should also be undertaken to determine the extent of any implicit priming associated with attending an identity parade (Wells, 1993) and the value of pre-parade instructions to combat this effect. There have already been several ventures into this area which have shown that the instructions given to witnesses prior to viewing a parade may be exerting an effect upon their likelihood of making an identification. For example, Rose, Bull & Vrji (2005) tested the effect of biased versus non-biased instructions on the rate of correct identifications and mistaken identifications in older adults. As a related measure, the authors asked each participant during debriefing to report what they remembered of the instructions given to them before viewing the parade. The findings showed that witnesses of all ages forgot being given the standard, non-biased instructions informing them that the culprit may or may not be present. Certainly more of the older witnesses forgot these instructions than the younger witnesses, which is explained in terms of age-related memory deficits, but most importantly, the witnesses who forgot these non-biased instructions made significantly more errors on both target present and target absent parades. In contrast, those witnesses who remembered the non-biased instructions made significantly fewer

errors, particularly when viewing a target absent parade. This is a finding that has been replicated to varying degrees in other papers (Brewer & Wells, 2006; Malpass & Devine, 1981; Steblay, 1997). Future research should attempt to address this issue in an attempt to resolve this systemic, implicit bias which leads eyewitnesses to make an identification.

Despite these considerations it is important to recognise that there are a substantial number of other factors which may affect eyewitness performance at an identity parade; the inherent difficulty of the task (Megreya & Burton, 2008; Kemp *et al.,* 1997; Valentine *et al.,* 2007; Megreya & Burton, 2006; Newell, Chiroro & Valentine, 1999), verbalisation effects (Meissner & Brigham, 2001; Schooler & Engstler-Schooler, 1990), change of appearance (Memon & Gabbert, 2003) memory (Ebbinghaus, 1885; Deffenbacher *et al.,* 2004; Ebbesen & Rienick, 1998; Kemp *et al.,* 2001; Bromby, 2002; Dunning & Stern, 1992; Christianson, 1992), situational factors such as weapon focus and stress (Steblay, 1992; MacLin, MacLin & Malpass, 2001; Maass & Kohnken, 1989; Valentine & Mesout, 2009; Morgan *et al.,* 2004; Deffenbacher *et al.,* 2004) and administrative issues (Wells *et al.,* 1998; Wells, 1993; Lindsay & Pozzulo, 1999; Steblay *et al.,* 2001; McQuiston-Surret *et al.,* 2006; Levi & Lindsay, 2001). It would be imprudent to declare one parade format superior purely on the facet of success rates when there are so many other influences to account for. On the basis of the current results V.I.P.E.R. appears to be fit for its purpose; reducing costs whilst preserving identification rates for both target present and target absent parades. V.I.P.E.R is certainly no poorer than the more traditional, simultaneous format.

The next stage of this thesis was to investigate the following step in an eye-witness's journey through the criminal justice system. Specifically what effect the use of 'Special Measures' exerts upon jurors perceptions of vulnerable adult witnesses; do they provoke a sympathetic response and boost credibility or are they reducing witness credibility and the likelihood of gaining a conviction? Crucially there are two criteria which are required to be met by any adult witness applying for permission to testify under these 'Special Measures': a proven history of intimidation in relation to their testimony and having made a positive identification

from a V.I.P.E.R. identity parade (Vital Voices: Helping Vulnerable Witnesses Give Evidence, 2003).

As discussed previously there is little literature which has tested the effects wrought upon jurors by observing a witness testifying using special measures (Goodman *et al.,* 1998; Swim *et al.,* 1993) and fewer still which have employed *adult* vulnerable witnesses. From an assessment of the currently published literature it appears that the current thesis may be the first to examine the effects that the special measures; use of a wooden screen, live CCTV link and pre-recorded video evidence, exert upon jurors' perceptions of vulnerable adult witnesses.

Certainly there has been some research examining the effects of live CCTV links for vulnerable child witnesses which has detailed a negative bias associated with testimony delivered in this manner. Goodman *et al.,* (1998) found that although there was no significant effect upon conviction rates there was a small but significant decline in the mock jurors' ratings of witness credibility despite perceiving them as more confident and consistent. Landstrom *et al.* (2005) reported that jurors who viewed live child witnesses reported them as being significantly more honest, more plausible and more confident than a child witness who testified via pre-recorded evidence. However, there are many other variables which jurors account for when judging child witnesses which may perhaps be a contributory factor of the decrease in credibility ratings (Quas, Thompson & Clarke-Stewart, 2005; Regan & Baker, 1998; Golding *et al.,* 2003; Wheatcroft *et al.,* 2004; Laimon & Poole, 2008).

There is also concern that the use of Special Measures during testimony may equally exert a negative effect upon jurors' perceptions of defendant guilt (Richards, Morris & Richards, 2008; Tausz & Ellison, 2005; Hoyano, 2001; Birch, 2000; Hamlyn *et al.,* 2006). There are claims that the use of Special Measures carries an implicit assumption of defendant guilt which may bias the jurors to the witness's favour and thus increase the possibility of a guilty conviction (Tausz & Ellison, 2005; Hoyano, 2001; Birch, 2000). The data from the jury

deliberations of Experiment 2 suggest that there is a small minority of the population who do form a link between the use of Special Measures and defendant guilt. In each of the three experimental conditions (Screen, CCTV and PVE) at least one mock juror raised this issue during deliberations and prompted a discussion of the validity of such a link. Equally, the majority of each jury group appeared to be aware that there are any number of factors which could have led to the witness electing to testify via a Special Measure. Of most interest thought to legal practitioners though is that, while there were small numbers of jurors who made a link between Special Measures and defendant guilt, this did not have any adverse effect upon the official delivered verdict. It did however make it more difficult to deliver a unanimous verdict; only the jurors who viewed the Screen condition delivered a unanimous verdict. Analysis of the audio recordings of the deliberations suggested that the jurors in the CCTV condition and the Pre-Recorded Video Evidence were at an impasse before being informed that a majority verdict would suffice. Whilst this is not such a significant effect in the United Kingdom it may lead to hung juries in jurisdictions require unanimity.

The data stemming from the initial series of experiments, where the jurors viewed the testimony and made judgements individually, suggested that the more removed the witness became from the courtroom the less credible they were rated by the mock jurors. There was no significant difference in witness ratings or conviction rates between the control, Open Court, condition and the Screen condition while there was a significant decline in witness credibility scores when they testified via a live CCTV link. However in a similar manner to other findings (Orcutt *et al.,* 2001; Swim *et al.,* 1993; Goodman *et al.,* 1998), there was no decrease in the rate of guilty verdicts. When the witness testified via Pre-recorded Video Evidence there was a significant decrease in witness credibility and believability ratings which was accompanied by a significant decline in conviction rates.

Such data was expected: there is a proven link that jurors perceive a lack of immediacy and emotional impact from testimony which is delivered from outside the trial courtroom (Davies & Noon, 1991) and that testimony vividness plays an important role in holding jurors attention and in their perception of witness credibility and plausibility (Nisbett & Ross, 1980;

Landstrom, Granhag & Hartwig, 2007; Colwell, Hiscock-Anisman, Memon, Rachel & Colwell, 2007). Indeed most of the data suggesting this theory has derived from CCTV link studies (Davies & Noon, 1991; Goodman *et al.,* 1998; Landstrom *et al.*, 2005, 2008; Orcutt *et al.*, 2001; Swim *et al.*, 1993) It is certainly not unreasonable to extend the prevailing theory to account for the significant decrease in witness credibility and conviction rates when testimony is delivered via pre-recorded video evidence. It is suggested that this perceived lack of immediacy, vividness and emotional impact originates from the jurors' inability to closely observe the witness's body language whilst testifying, a significant factor widely used by jurors to determine deceit and credibility (Aron & Rosner, 1998; Boccaccini, 2002; Mehrabian, 1981; Landry & Brigham, 1992). It is further suggested that as the witness becomes increasingly distal from the jurors the weight given to their testimony diminished as the jurors pay less attention to it (Nisbett & Ross, 1980) and that physical proximity to the witness allows the jurors to identify with them which results in more favourable ratings of credibility and plausibility (Landstrom *et al.,* 2005, 2007; Goodman *et al.*, 1998; Orcutt *et al.*, 2001). It is easy to understand this association as both the CCTV link and pre-recorded video evidence allow only the witness's head and shoulders to be seen whereas more gestures and behaviour can be observed live in-court. Pre-recorded Video Evidence then takes a further step in diminishing physical proximity, vividness and emotional impact: the testimony is often recorded weeks and months in advance and there is no need to recount traumatic experiences in a public courtroom and so the witness is typically less anxious. It may be that this decreased anxiety is construed as a lack of emotion, and then, in turn, an indicator of dishonesty (Golding *et al.,* 2003; Wheatcroft *et al.*, 2004; Regan & Baker, 1998; Boccaccini, 2002; Aron & Rosner, 1998; Landstrom *et al.*, 2008).

The current data clearly illustrates that there is a progressive decline in the ratings of witness credibility, believability and rates of conviction as the proximity between witness and jurors increases; a finding that supports much of the current literature in this area (Goodman *et al.*, 1998, 2004; Orcutt *et al.,* 2001; Swim *et al.,* 1993; Landstrom *et al.*, 2005, 2005, 2008). However, while these results are of some value there is an intrinsic limitation to the initial experimental series; jurors do not deliberate in isolation but serve as part of a larger group to

deliver a collective judgement upon the case. Thus Experiment 2 was designed to determine if the effects of the Special Measures upon jurors' perceptions were replicated when a deliberation stage was introduced.

The first analysis examined if there was a significant effect of deliberation; did the jurors' ratings of the witness and likelihood of conviction significantly change after deliberation with the other jurors. The data is clear on this point; there is no evidence to suggest that the individual juror's perceptions were altered significantly by the group deliberations. There was evidence of a small decreasing trend but this was not a significant effect. This finding supports the hypothesis that initial juror perceptions, once formed, are firmly set and highly unlikely to be significantly altered (Stasser *et al.,* 1982; McKelvey & Kerr, 1988; Asch, 1951).

The second analysis was performed entirely on the 'after deliberation' data to determine whether there were any significant differences between the four trial conditions when presented to a deliberating group. There was a general increasing trend with ratings of the witness being lowest in the Open Court condition and rising steadily with decreasing witness proximity. However, the resultant analysis of the data revealed that there was no significant effect of testimony presentation method upon the mean ratings of vulnerable witness credibility and believability. There was a similar pattern within the verdict data; there is a general increasing trend as the jurors' proximity to the witness decreases when we look at the Open Court (14.8%), Screen (15.38%) and CCTV link (24.8%) data. Yet, when Pre-recorded Video Evidence was used the rate of guilty verdicts fell back to 15.4%. While this pattern is readily evident, not one of the three experimental conditions differed significantly from the control condition.

We can plainly see that the 'after deliberation' data elicited an opposite effect to that seen in the first experimental series. When the jurors view the testimony and deliberate as a group there is an obvious increase in the mean ratings of witness credibility, believability and the number of guilty verdicts as the Special Measures increasingly isolates the witness from the jurors. It was speculated that this opposite trend derives from one, or more, of the 13

member jury group proposing the use of a Special Measure as an implicit suggestion of defendant guilt which is developed into increased witness confidence. However, whilst there is evidence from the audio recordings that there were a small number of jurors in each experimental jury group that proposed this link between defendant guilt and use of the Special Measure, there is no significant evidence that this influenced the jurors in the witness' favour and led to an increased likelihood of a guilty verdict. While this is reassuring for legal practitioners, it does not account for the current findings. It is alternatively suggested that the mock jurors may have believed the defendant to be culpable but equally felt that the prosecution had failed to satisfy the burden of proof. Indeed evidence from the deliberation recordings show that at least one mock juror suggested that they would prefer to return a 'not proven' verdict for just such a reason. Whilst Scots Law does have this third verdict option and despite the fact that the mock trials were run in accordance with Scots Law, the jurors were limited to either a 'guilty' or 'not guilty' verdict decision. This steps were taken as many of the participants recruited were not Scottish and were considered to be unfamiliar with the 'not proven' verdict and its interpretation but also because there has been little research undertaken of this option. The extant literature on this topic has revealed that there is a poor understanding of the 'not proven' verdict (Hope *et al.*, 2008). Additionally, the 'not proven' verdict option typically accounts for only one-fifth of acquittals in Scotland (Scottish Office Study, 2004) and is associated with much controversy (http://www.siliconglen.com/Scotland/1_8.html; Bray, 2005; Duff, 1999; Davies, 1996; Barbato, 2005; http://news.bbc.co.uk/1/hi/scotland/6500541.stm). As the 'not proven' verdict results in the same outcome as a 'not guilty' verdict, acquittal, it was decided to use only 'not guilty' as a means of simplifying the methodology.

While there were no significant differences in testimony presentation method identified, the trend is unmistakable and therefore the data raises some serious questions for the judicial system. While it is certainly advantageous to vulnerable witnesses to be permitted to testify using a variety of Special Measures, the data from this thesis provides some small corroborating evidence for claims of an implicit negative bias of defendant guilt (Tausz & Ellison, 2005; Hoyano, 2001; Birch, 2000).

In summary, the data from this thesis provides both a valuable start point for further investigations in this particular area and, more importantly, provides quantitative evidence as to the success of the provisions of The Vulnerable Witnesses Act. There is weak evidence that when Special Measures are used in court they carry an implicit connotation of defendant guilt before testimony is even begun and this prompts the mock jurors to rate the vulnerable adult witness as more credible over the witness who testified in an open court. However, this finding was not directly tested by the current study and further research is necessary to firmly establish this implicit link. There is a clear increasing trend in mean witness ratings as the Special Measures invoked increasingly remove the witness from the scrutiny of the jurors. This is in contrast to the extant literature which claims that as proximity to the witness decreases, the ratings of witness credibility decrease due to a decline in testimony vividness (Nisbett & Ross, 1980; Colwell *et al.*, 2007; Landstrom *et al.*, 2007), emotional impact (Davies & Noon, 1991) and immediacy (Nisbett & Ross, 1980; Davies & Noon, 1991). However, all of these studies have used children as their vulnerable witness whilst the current research used a vulnerable adult witness. To date, this is the only paper which has tested mock jurors' perceptions of adult witnesses who elect to testify via Special Measures. Thus, it may be that these findings are typical and so further investigations are required to corroborate the data and clearly define the effects of Special Measures on jurors' perceptions of vulnerable adult witnesses.

Not one of the experimental conditions elicited a significant increase in the number of guilty verdicts over that associated with the control condition, Open Court. Therefore it is claimed that despite the increase in perceived witness credibility, and thus negative bias towards the defendant, there is no evidence to suggest that the use of Special Measures erodes the judicial rights of the defendant. There is a persuasive suggestion that the vulnerable witness's testimony is delivered more effectively and that the witness experience of trial and testimony is less distressing (Richards, Morris & Richards, 2008; Birch, 2000) when Special Measures are invoked.

However, there appears to be an extraneous effect of using Special Measures upon witnesses' later reflections; Richards, Morris & Richards (2008) reported that vulnerable witnesses stated after the conclusion of their trial that they would not be willing to testify in any other trial without the same considerations, a finding also reported by Birch in her 2000 review of the Special Measures provisions in England and Wales. This means that they would refuse to testify if they were not permitted to use Special Measures, even though they may not meet the stringent criteria in a different trial. It is therefore suggested that the witness experience in general be improved. For example, many courts still use communal waiting rooms for both prosecution and defence witnesses; the issues regarding this are evident, exposing witnesses to further conflict and stress. Furthermore, the Richards, Morris & Richards (2008) survey found that witnesses feel isolated from both the route to trial, judgements and sentencing, often claiming that they hear nothing between attending an identity parade and receiving a citation to appear at court. This lack of information leads to the witness feeling unvalued by the judicial system and contributes to their later reluctance to testify in other trials (Hamlyn *et al.,* 2006; Birch, 2000). It seems that improving the witness's sense of personal significance to the trial (Burton, Evans & Sanders, 2006; Richards, Morris & Richards, 2008) may go some way towards ameliorating the pervading negative view of the witness experience at court. It may be beneficial to the judicial system as a whole that the procedures of the court be further modified in line with these recommendations to improve the *general* experience of witnesses summoned to testify rather than focusing on a small sub-population which may be exacerbating reluctance to testify.

There are of course several limitations of the Special Measures experiments. The extent to which the mock jurors understood the 'vulnerability' of the adult witness is undetermined. Under the terms of The Vulnerable Witnesses (Scotland) Act 2003, jurors are not told the reasons behind the judge's decision to permit Special Measures during testimony. This is obviously to prevent any bias being sparked in the jurors' perception of the defendant and to allow the defendant a fair trial. Future research should aim to investigate these instructions to the jury more closely; would specific instructions to the jury reduce the influence that the

use of Special Measures exerts over juror perceptions of witnesses? It would also be advantageous to determine exactly how much the public understands about the provisions for vulnerable witnesses; what characterises them as vulnerable? Experiment 2 provided valuable data regarding the deliberations of mock jury groups although no significant difference were found. Yet several factors were very close to the level of significance which may be a consequence of the small sample size allocated to each of the jury groups (n = 13). Future research should expand these sample sizes by running several different jury groups for each of the four conditions. Finally, further research is required to clearly identify the cues used by mock jurors when making their judgments of witness credibility and believability. This could take the form of either open-ended questions during debriefing or through including self-report options on the measurement questionnaires. This research should help to determine exactly how jurors perceive vulnerable adult witnesses and what behavioural, verbal and emotional cues lead them to these perceptions. Finally, the use of jury trials is decreasing every year (Tinsley, 2001) with the rise of civil trials presided over by a single judge or sheriff. The data generated from Experiments 1(a) – 1(c) provide some data as to the effects that Special Measures exert upon a single arbiter; the use of Special Measures during testimony was associated with a decrease in ratings of witness credibility and a decrease in the number of convictions, particularly when Pre-recorded Video Evidence was employed. While these data give us an early impression of the effects Special Measures exert upon a single arbiter in a trial it must be noted that the participants in these experiments were predominantly students and none were trained legal professionals. Future research should attempt to examine the effect of Special Measures exert upon judges and sheriffs perceptions of vulnerable adult witnesses.

In conclusion, there remain several obstacles on the eye-witness's route through the criminal justice system. While there is no doubt that the introduction of V.I.P.E.R. has significantly minimised witness apprehension (Slater, 1995; Steblay *et al.,* 2001) and therefore caused a large reduction in police costs whilst improving efficiency (McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Lindsay & Pozzulo, 1999; Levi & Lindsay, 2001; Slater, 1995) it is apparent that there is no evidence of a significant superiority for the more problematic target

absent parades. It is accepted that these findings are due to the limitations enforced under PACE Code D regulations and are unlikely to be altered radically however, in its current hybrid format there will continue to be a large percentage of misidentifications although, fortunately, these should no longer progress to trial (Valentine & Heaton, 1999; West Yorkshire Police, 2009). It is recommended that V.I.P.E.R. move progressively towards a wholly sequential format using the procedural guidelines as proposed widely in the literature (McQuiston-Surrett *et al.,* 2006; Steblay *et al.,* 2001; Lindsay & Pozzulo, 1999; Levi & Lindsay, 2001; Kassin *et al.,* 2001). In the interim period, as discussed earlier, further examination of V.I.P.E.R. under its current and proposed formats should be actively pursued.

On progressing to trial it appears that the wishes and needs of the vulnerable adult witness are being suitably accommodated by the introduction of several Special Measures (Birch, 2000; Richards, Morris & Richards, 2008; Hamlyn *et al.,* 2004). Witness groups report a substantial decrease in witness anxiety and witnesses themselves declare themselves to be largely happy with the provisions (Birch, 2000; Hamlyn *et al.,* 2004; Richards, Morris & Richards, 2008). The data suggests that it is defendants, solicitors and judges who should be most concerned. There is some small evidence that the use of Special Measures does carry an implicit negative bias leading to biased jurors perceptions of both witness and defendant. Although there was no significant negative impact upon the number of guilty verdicts in the current study it is strongly recommended that further investigations be conducted before a definitive conclusion is proposed. At this time though it is proposed that policymakers require that all jury groups be given specific instructions regarding their perception of both witness and defendant in cases where Special Measures have been permitted, *before* proceeding to testimony. It may be beneficial to explicitly define all the situations and criteria which could lead to Special Measures being used by a vulnerable witness. By improving jurors' understanding of the provisions of The Vulnerable Witnesses (Scotland) Act (2003) the implicit link between the use of Special Measures and defendant guilt, evidenced in the deliberations of the current study, may be reduced.

The data generated from this thesis provides a valuable and important starting point for further and more detailed investigation of both eyewitness identification methods and the use of Special Measures in court. While the conclusions drawn are a reasonable account of the data it is prudent to pursue further investigations in order to form a coherent theory for the effects discussed within. It is the opinion of the author that the recent legislative steps of V.I.P.E.R. and Special Measures are certainly advantageous and a positive step for eye-witnesses, yet, as discussed, there remains room for further improvement for the benefit of witnesses, defendants and the criminal justice system as a whole.

References

Ahola, A. S., Christianson, S. A. & Hellstrom, A. (2009) Justice Needs a Blindfold: Effects of Gender and Attractiveness on Prison Sentences and Attributions of Personal Characteristics in a Judicial Process. *Psychiatry, Psychology and Law, 16*(S1), S90-S100.

Aron, R. & Rosner, J.L. (1998) *How to Prepare Witnesses for Trial.* 2nd Edition, West:Danvers, MA.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow (ed.) *Groups, leadership and men*. Pittsburgh, PA: Carnegie Press.

Bell, B. E., & Loftus, E. (1985). Vivid persuasion in the courtroom. *Journal of Personality Assessment*, *49*, 659–664.

Birch, D. J. (2000) A Better Deal for Vulnerable Witnesses. *Criminal Law Review,* April, 223-249.
Birch, S. A. J., Akmal, N., & Frampton, K. L. (2010) Two-year-olds are vigilant of others' non-verbal cues to Credibility. *Developmental Science, 13*(2), 363-369.

Blamey, R. K., McCarthy, P. & Smith, R. (2000) Citizens' Juries and Small-group Decision-making. Retrieved on 09/05/2009 from http://cjp.anu.edu.au/docs/CJ2.pdf

Boccaccini, M.T. (2002) What do we really know about witness preparation? *Behavioural Sciences and the Law, 20,* 161 – 189.

Boccaccini, M. T. & Brodsky, S. L. (2002) Believability of Expert and Lay Witnesses: Implications for Trial Consultation. *Professional Psychology: Research and Practice, 33*(4), 384-388.

Boccaccini, M. T., Gordon, T. & Brodsky, S. L. (2005) Witness Preparation Trainingwith Real and Simulated Criminal Defendants. *Behavioural Sciences and the Law, 23*, 659-387.

Bornstein, B. H. (1999) The Ecological Validity of Jury Simulations: Is the Jury Still Out?, *Law and Human Behavior, 23*, (1), 75-91.

Bradfield, A. & Wells, G. L. (2005) Not the same old hindsight bias: Outcome information distorts a broad range of retrospective judgments. *Memory & Cognition, 33*(1), 120-130.

Bray, R. M., & Kerr, N. L. (1982). Methodological considerations in the study of the psychology of the courtroom. In N. L. Kerr & R. M. Bray (Eds.), *The psychology of the courtroom* (pp. 287–323). New York: Academic Press.

Brewer, N. & Palmer, M. A. (2010) Eyewitness identification tests. *Legal and Criminological Psychology, 15,* 77-96.

Brigham, J. C. & Barkowitz, P. (1978) Do "They all look alike"? The effect of race, sex, experience, and attitudes on the ability to recognize faces. *Journal of Applied Social Psychology, 8*(4), 306-318.

Bromby, M. (2002) To be taken at face value? Computerised Identification. *Informations & Communications Technology Law, 11*(1), 63-73.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001) Matching Identities of Familiar and Unfamiliar faces Caught on CCTV Images. *Journal of Experimental Psychology: Applied, 7*(2), 207-218.

Bruce, V., Valentine, T. & Baddeley, A. (1987). The basis of the three-quarters view advantage in face recognition. *Applied Cognitive Psychology*, *1*, 109-120.

Bruine de Bruin, W. & Keren, G. (2003) Order effects in sequentially judged options due to the direction of comparison. *Organizational Behaviour and Human Decision Processes, 92,* 91 – 101.

Bruine de Bruin, W. (2005) Save the last dance for me: unwanted serial position effects in jury evaluations. *Acta Psychologica, 118,* 245 – 260.

Burton, M., Evans, R., and Sanders, A. (2006) Are special measures for vulnerable and intimidated witnesses working? Evidence from the criminal justice agencies. Home Office Online Report 01/06

Burton, A.M, Wilson, S., Cowan, M & Bruce, V. (1999). Face recognition in poor quality video: evidence from security surveillance. *Psychological Science, 10*, 243-248.

Busey, T. A. & Loftus, G. A. (2007) Cognitive science and the law. *Trends in Cognitive Science, 11*(3), 111-117.

Carlson, K. A., Gronlund, S. D. & Clark, S. E. (2008) Lineup Composition, Suspect Position, and the Sequential Lineup Advantage. *Journal of Experimental Psychology:Applied, 14*(2), 118-128.

Carlson, K.A. & Russo, E.J. (2001) Biased Interpretation of Evidence by Mock Jurors. *Journal of Experimental Psychology: Applied, 7,* 91 – 103.

Casper, J., Benedict, K. & Perry, J. (1989). Juror decisionmaking, attitudes, and the hindsight bias. *Law and Human Behavior, 13*, 291-310.

Castellow, W. A., Wuensch, K. L. & Moore, C. L. (1990) Effects of physical attractiveness of the plaintiff and defendant in sexual harassment judgments. *Journal of Social Behavior & Personality, 5(6),* 547-562.

Charman, S. D., & Wells, G. L. (2008). Can eyewitnesses correct for external influences on their lineup identifications? The actual/counterfactual assessment paradigm. *Journal of Experimental Psychology: Applied, 14*(1), 5-20.

Christianson, S. A. (1992) Emotional Stress & eyewitness memory. A critical review. *Psychological Bulletin, 112*(2), 284-309.

Clark, S. E. & Davey, S. L. (2005) The Target-to-Foils Shift in Simultaneous and Sequential Lineups. *Law and Human Behavior, 29*(2), 151-172.

Clark, S. E., Marshall, T. E. & Rosenthal, R. (2009) Lineup Administrator Influences on Eyewitness Identification Decisions. *Journal of Experimental Psychology: Applied, 15*(1), 63-75.

Clark, S. E. & Tunnicliff, J. L. (2001) Selecting Lineup Foils in Eyewitness Identification Experiments: Experimental Control and Real-World Simulation. *Law and Human Behavior, 25*(3), 199-216.

Cliff, L. & Burton, A. M. (2006) Witness Confidence and Presentation Order: A New Route to Juror Manipulation? Unpublished Manuscript.

Colwell, K., Hiscock-Anisman, C., Memon, A., Rachel, A. & Colwell, L. (2007) Vividness and Spontaneity of Statement Details Characteristics as Predictors of Witness Credibility. *American Journal of Forensic Psychology, 25*(1), 5-30.

Conley JM, O'Barr, Lind EA. 1978. The power of language: presentation style in the courtroom. *Duke Law Journal 1978*(6): 1375–1399.

Cross, J.F., Cross, J., & Daly, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. Perception and Psychophysics, 10, 393–396.

Daftary-Kapur, T., Dumas, R. & Penrod, S. D. (2010) Jury decision-making biases and methods to counter them. *Legal and Criminological Psychology, 15*(1), 133-154.

Darby, B. W. & Jeffers, D. (1988) The Effects of Defendant and Juror Attractiveness on Simulated Courtroom Trial Decisions. *Social Behaviour and Personality: An International Journal, 16*(1), 39-50.

Davies, G. & Noon, E., (1991). An Evaluation of the Live Link for Child Witnesses. London: Home Office.

Davis, J.H., Kerr, N.L., Stasser, G., Meek, D., & Holt, R. Victim consequences, sentence severity, and decision processes in mock juries. *Organizational Behavior and Human Performance*, *18*, 346–365.

Davies, G. & Thasen, S. (2000). Closed-circuit television: How effective an identification aid? *British Journal of Psychology, 91,* 411-426.

Davies, G.M., Wilson, J.C., Mitchell, R. & Milsom, J. (1995). Videotaping children's evidence: an evaluation. London: HMSO.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., and McGorty, E. K. (2004) A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory. *Law and Human Behavior, 28*(6), 687-706.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., and McGorty, E. K. (2008) Forgetting the Once-Seen Face: Estimating the Strength of an Eye-witness's Memory Representation. *Journal of Experimental Psychology: Applied, 14*(2), 139-150.

Devine, D.J., Clayton, L.D., Dunford, B.B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law*, *7*, 622-727.

Diamond, S. S. (1997). Illuminations and shadows from jury simulations. *Law and Human Behavior, 21*, 561–571.

Dixon, S. & Memon, A. The Effect of Post-Identification Feedback on the Recall of Crime and Perpetrator Details. *Applied Cognitive Psychology, 19*, 935-951.

Douglass, A. B. & McQuiston-Surrett, D. (2006) Post-identification feedback: exploring the effects of sequential photospreads and eyewitnesses' awareness of the identification task. *Applied Cognitive Psychology, 20*(8), 991-1007.

Douglass, A. B. & Steblay, N. (2006) Memory distortion in eyewitnesses: a meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology, 20*(7), 850-869.

Dunning, D. & Stern, L. B. (1992) Examining the generality of eyewitness hypermnesia: A close look at time delay and question type. *Applied Cognitive Psychology, 6*(7), 643-657.

Ebbesen, E. B., & Flowe, H. D. (2002) The effect of lineup member similarity on recognition accuracy in simultaneous and sequential lineups. *Law and Human Behavior, 31*(1), 33-52.

Ebbesen, E. B., & Rienick, C. B. (1998) Retention Interval and Eyewitness Memory for Events and Personal Identifying Attributes. *Journal of Applied Psychology, 83*(5), 745-762.

Ebbinghaus, H. (1885) Memory: A Contribution to Experimental Psychology. Translated by H. A. Ruger & C. E. Bussenius (1913). New York: Teachers College, Columbia University.

Editorial (2004) Vulnerable and Intimidated Witnesses: Are Special Measures Working? *Criminal Law Review,* 599-600.

Elliott, R. (1998). Vulnerable and Intimidated Witnesses: A Review of the Literature. In *Speaking Up for Justice*. London: Home Office. Eyewitness Misidentifications.

Erickson B, Lind E. A, Johnson B.C., O'Barr W. M. (1978) Speech style and impression formulation in a court setting: the effects of 'powerful' and 'powerless' speech. *Journal of Experimental Social Psychology 14,* 266–279.

The Innocence Project, Eyewitness Misidentifications. Retrieved 25 May 2009 from http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php

Ferguson, D. P., Rhodes, G., Lee, K., & Sriram, N. (2001). "They all look alike to me": Prejudice and cross-race face recognition. *British Journal of Psychology, 92*, 567-577.

Findlay, M. (1994) *Jury Management in New South Wales.* Australian Institute of Judicial Administration.

Finkel, N. J., & Duff, K. B. (1991). Felony-murder and community sentiment: Testing the Supreme Court's assertions. *Law and Human Behavior, 15,* 405–429.

Finkel, N. J., & Handel, S. E (1989). How jurors construe "insanity." *Law and Human Behavior, 13,* 41–59.

Fulero, S. M., & Finkel, N. J. (1991). Barring ultimate issue testimony. *Law and Human Behavior, 15*, 495–507.

Fusaro, M. & Harris, P. L. (2008) Children assess informant reliability using bystanders' non-verbal cues. *Developmental Science, 11*(5), 771-777.

Fyfe, N. R. & McKay, H. (1998) Desperately Seeking Safety. *The British Journal of Criminology, 40*(4), 675-691.

Golding, J. M., Fryman, H. M., Marsil, D. M., & Yozwiak, J. A. (2003). Big girls don't cry: The effect of child witness demeanor on juror decisions in a child sexual abuse trial. *Child Abuse and Neglect,* 27*,* 1311-1321*.*

Golding, J. M., Sego, S. A., Polley, R. R., & Hasemann, D. (1995). The believability of repressed memories. *Law and Human Behavior, 19,* 569-592.

Goldstein, A. G. & Chance, J. E. (1979) Do "foreign" faces really look alike? *Bulletin of the Psychonomic Society. 13*(2), 111-113.

Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., Orcutt, H., Thomas, S., Shapiro, C. & Sachsenmaier, T. (1998) Face-to-face confrontation: effects of closed-circuit technology on children's eyewitness testimony and jurors' decisions. *Law and Human Behavior, 22*, (2), 165-203.

Gronlund, S. D. (2004) Sequential lineups: shift in criterion or decision strategy? *Journal of Applied Psychology, 89*, 362-368.

Hamlyn, B., A. Phelps, J. Turtle and G. Sattar (2004). Are Special Measures Working? Evidence From Surveys of Vulnerable and Intimidated Witnesses. Home Office Research Study, No. 283. London: Home Office Research, Development and Statistics Directorate.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000) Recognition of unfamiliar faces. *Trends in Cognitive Science, 4*(9), 330-337.

Hancock, K.J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology, 99*, 45–56.

Hastie, R., Penrod, S., & Pennington, N. (1983). What goes on in a jury deliberation? *American Bar Association Journal*, *69*, 1848-1853.

Henderson, Z., Bruce, V. & Burton, A.M. (2001). Matching the faces of robbers captured on video. *Applied Cognitive Psychology, 15,* 445-464.

Hope, L., Greene, E., Memon, A., Gavisk, M. & Houston, K. (2008) A Third Verdict Option: Exploring the Impact of the Not Proven Verdict on Mock Juror Decision Making. *Law & Human Behaviour, 32,* 241-252.

House of Commons, Parliamentary Publications. Retrieved 13 April 2007 from http://www.publications.parliament.uk/pa/cm200506/cmstand/deleg1/st060508/60508s01.htm

Hoyano, L. C. H. (2001) Striking a balance between the rights of defendants and vulnerable witnesses: Will Special Measures Directions contravene guarantees of a fair trial? *Criminal Law Review,* December, 948-969.

Johnson, K. J. & Fredrickson, B. L. (2005) "We All Look the Same to Me" Positive Emotions Eliminate the Own-Race Bias in Face Recognition. *Psychological Science, 16*(11), 875-881.

Kassin, S. M., Tubb, A., Hosch, H. M., & Memon, A. (2001). On the 'General Acceptance' of eyewitness testimony research: A new survey of experts. *American Psychologist, 56*, 405-416.

Kassin, S. M. & Wrightsman, L. S. (1979) On the requirements of proof: The timing of judicial instruction and mock juror verdicts. *Journal of Personality and Social Psychology, 37*(10), 1877-1887.

Kebbell, M.R., & Giles, D.C. (2000). Some experimental influences of lawyers' complicated questions on eyewitness confidence and accuracy. *Journal of Psychology, 134*, 129-139.

Kemp, R. I., Pike, G. E. and Brace, N. A. (2001). Video-based identification procedures: combining best practice and practical requirements when designing identification systems. *Psychology, Public Policy, and Law*, *7*(4), pp. 802–807.

Kemp, R., Towell, N. & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology, 11*, 211-222.

Kneller, W., Memon, A., & Stevenage, S. (2001) Simultaneous and Sequential line-ups: Decision process of accurate and inaccurate witnesses. *Applied Cognitive Psychology, 15*, 659-671.

Kramer, T.H., Buckhout, R., & Eugenio, P. (1990). Weapons focus, arousal, and eyewitness memory: Attention must be paid. *Law and Human Behavior, 14*, 167-184.

Kramer, G. P., Kerr, N. L. & Carroll, J. S. (1990) Pre-trial publicity, judicial remedies, and jury bias. *Law and Human Behavior, 14,* (5), 409-438.

Laimon, R. L. & Poole, D. A. (2008) Adults Usually Believe Young Children: The Influence of Eliciting Questions and Suggestibility Presentations on Perceptions of Children's Disclosures. *Law & Human Behaviour, 32,* 489-501.

Lampinen, J. M., Scott, J. N., Pratt, D., Leding, J. K., & Arnal, J. D. (2007). The effects of warning on the post-identification feedback effect. *Applied Cognitive Psychology, 21*, 1037-1056.

Landry, K. & Brigham, J. C. (1992), The effect of training in criteria-based content analysis on the ability to detect deception in adults. *Law and Human Behavior*, Vol. 16 pp.663 - 675.

Landström, S., Granhag, P.A., & Hartwig, M. (2005) Witnesses appearing live versus on video: Effects on Observer's Perception, Veracity Assessments and Memory. *Applied Cognitive Psychology, 19*, 913-933.

Landström, S., Granhag, P.A., & Hartwig, M. (2007) Children's live and videotaped testimonies: How presentation mode affects observers' perception, assessment and memory. *Legal and Criminological Psychology, 12,* 333-347.

Levi, A. M., & Lindsay, R. C. L. (2001) Lineup and photo spread procedures: Issues concerning policy recommendations. *Psychology, Public Policy, and Law, 7*, 776-790.

Levi, A. M. (1998) Are defendants guilty if they were chosen in a lineup? *Law and Human Behavior, 22*, 389-407.

Lind E. A., Erickson, B., Conley J. M., & O'Barr, W. M. (1978) Social attributions and conversational style in trial testimony. *Journal of Personality and Social Psychology 36*, 1558–1567.

Lindsay, R.C.L., Lea, J., & Fulford, J. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology*, *76*, 741-745.

Lindsay, R. C. L., & Pozzulo, J. D. (1999) Sources of Eyewitness Identification Error. *International Journal of Law and Psychiatry, 22*(3-4), 347-360.

Lindsay, R. C. L., & Wells, G. L. (1985) Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*, 556-564.

Liu, C. H., & Chaudhuri, A. (2001) Re-assessing the three quarter face effect in face recognition. *Cognition, 83*(1), 31-48.

Liu, C. H., Seetzen, H., Burton A. M., & Chaudhuri, A. (2003) Face Recognition Is Robust With Incongruent Image Resolution: Relationship to Security Video Images. *Journal of Experimental Psychology: Applied, 9*(1), 33-41.

Lipton, J. P. (1977) On the psychology of eyewitness testimony. *Journal of Applied Psychology, 66*, 79-89.

Loftus, E.F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560-572.

Luus, C. A. E. & Wells, G. L. (1994) The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology, 79*(5), 714-723.

Luus, C. A. E. & Wells, G. L. (1991) Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15*(1), 43-57.

Maas, A. & Kohnken, G. (1989) Eyewitness Identification: Simulating the "Weapon Effect". *Law and Human Behavior, 13*(4), 397-408.

MacLin, O. H., MacLin, M. K, & Malpass, R. S. (2001) Race, arousal, attention, exposure and delay: An examination of factors mediating face recognition. *Psychology, Public Policy and Law, 7*(1), 134-152.

Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law & Human Behavior, 5*, 299-309.

Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482-489.

Malpass, R. S., & Lindsay, R. C. L. (1999). Measuring lineup fairness. *Applied Cognitive Psychology, 13*, S1-S7.

McAuliff, B. D. & Kovera, M. B. (2007) Estimating the Effects of Misleading Information on Witness Accuracy: Can Experts Tell Jurors Something They Don't Already Know? *Applied Cognitive Psychology, 21*, 849-870.

McCabe, S. and Purves, R. (1974) *The Shadow Jury at Work.* Oxford University Penal Research Unit: Blackwell.

McCoy M. L., Nunez N. & Dammeyer M. M. (1999) The Effect of Jury Deliberations on Jurors' Reasoning Skills. *Law and Human Behavior, 23,* (5), 557-575.

McKelvey, W. & Kerr, N. H. (1988). Differences in conformity among friends and strangers. *Psychological Reports*, *62*, 759-762.

McQuiston-Surrett, D., Malpass, R. S., & Tredoux, C. G. (2006) Sequential vs Simultaneous Lineups: A Review of Methods, Data and Theory. *Psychology, Public Policy, and Law, 12*(2), 137-169.

Mecklenburg, S. H. (2006) Report to the Legislature of the State of Illinois: The Illinois Pilot Program on Sequential Double-Blind Identification Procedures.

Megreya, A. M., & Burton, A. M. (2008) Matching Faces to Photographs: Poor Performance in Eyewitness Memory (Without the Memory). *Journal of Experimental Psychology: Applied, 14*(4), 364-372.

Megreya, A. M. & Burton, A. M. (2006) Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition, 34,* (4), 865-876.

Mehrabian, A. (1981). *Silent messages: Implicit communication of emotions and attitudes* (2nd Ed.). Belmont, CA: Wadsworth.

Meissner, C. A., & Brigham, J. C. (2001) A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology, 15*(6), 603-616.

Meissner, C.A., & Brigham, J.C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy and Law, 7,* 3–35.

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005) Eyewitness decisions in simultaneous and sequential lineups: A dual process signal detection theory analysis. *Memory & Cognition, 33,* 783-792.

Memon, A & Gabbert, F. (2003). Unravelling the effects of a sequential lineup. *Applied Cognitive Psychology, 6*, 703-714

Memon, A., Hope, L. & Bull, R. (2003) Exposure Duration: Effects of Eye-witness confidence and accuracy. *British Journal of Psychology, 94*, 339-354.

Moreland, R. L., Levine, J. M., & Wingert, M. L. (1996). Creating the ideal group: Composition effects at work. In J. Davis & E. Witte (Eds.), *Understanding group behavior (Vol. 2, pp. 11-35).* Mahwah, NJ: Erlbaum.

Morgan III, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M. & Southwick, S. M. (2004) Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry, 27*(3), 265-279.

Nemeth, C. (1981). Jury trials: Psychology and the law. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 14, pp. 309-367). New York: Academic Press

Newell, F. N., Chiroro, P. & Valentine, T. (1999). Recognising unfamiliar faces: The effects of distinctiveness and view. *Quarterly Journal of Experimental Psychology*, 52(A), 509-534.

Nisbett, R., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs, NJ: Prentice-Hall.

O'Barr, W. (1974) cited in Boccaccini, M.T. (2002) What do we really know about witness preparation? *Behavioural Sciences and the Law, 20*, 161 – 189.

O'Barr, W. M. (1982) *Linguistic Evidence: Language, Power, and Strategy in the Courtroom*. Academic: New York.

O'Barr W.M., & Conley J.M. (1976) When a juror watches a lawyer. *Barrister 3*(2): 8–11, 33.

Orcutt, H. K., Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., & Thomas, S. (2001) Detecting deception in children's testimony: factfinders' abilities to reach the truth in open court and closed-circuit trials. *Law and human behavior*, *25*(4):339-72.

Pedzek, K., Blandon-Gitlen, I. & Moore, C. (2003) Children's face recognition memory: More evidence for the cross-race effect. *Journal of Applied Psychology, 88*(4), 760-763.

Pennington, N. & Hastie, R. (1992) Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of Personality and Social Psychology. 62*(2), 189-206.

Penrod, S. & Cutler, B. (1995) Witness confidence and witness accuracy: Assessing their Forensic Relation. *Psychology, Public Policy and Law, 1,* 817 – 845.

Pike, G., Kemp, R., Brace, N., Allen, J. & Rowlands, G. (2000) The effectiveness of video identification parades. *Proceedings of The British Psychological Society, 8*(1), 44.

Porter, S., Doucette, N. L., Woodworth, M., Earle, J. & MacNeil, B. (2008) Halfe the world knowes not how the other halfe lies: Investigation of verbal and non-verbal signs of deception exhibited by criminal offenders and non-offenders. *Legal and Criminological Psychology, 13*, 27-38.

Pozzulo, J. D., Crescini, C., & Panton, T. (2008) Does methodology matter in eyewitness identification research?: The effect of live versus video exposure on eyewitness identification accuracy. *International Journal of Law and Psychiatry, 31*(5), 430-437.

Quas, J. A., Thompson, W. C., & Clarke-Stewart, A. (2005). Do jurors know what isn't so about child witnesses? Law and Human Behavior, 29, 425-456.

Read, D. J., Vokey, J. R. & Hammersley, R. (1990) Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(5), 870-882.

Regan, P. C., & Baker, S. J. (1998). The impact of child witness demeanor on perceived credibility and trial outcome in sexual abuse cases. *Journal of Family Violence*, 13, 187-195.

Richards, P., Morris, S. & Richards, E. (2008) Turning up the Volume: The Vulnerable Witnesses (Scotland) Act 2004. Scottish Government Social Research.

Rose, R. A., Bull, R. & Vrij, A. (2005) Non-biased lineup instructions do matter - a problem for older witnesses. *Psychology, Crime and Law, 11*(2), 147-159.

Ross, D. F., Hopkins, S., Hanson, E., Lindsay, R. C. L., Hazan, K., & Eslinger, T. (1994). The impact of protective shields and videotape testimony on conviction rates in simulated trial of sexual abuse. *Law and Human Behavior, 18*, 553–566.

Ross, S. J. & Malpass, R. S. (2008) Moving Forward: Response to ''Studying Eyewitness Investigations in the Field''. *Law and Human Behaviour, 32,* 16-21.

Schacter, D. L., Dawes, R., Jacoby, L. L., Kahneman, D., Lempert, R., Roediger, H. L. & Rosenthal, R. (2008) Policy Forum: Studying Eyewitness Investigations in the Field. *Law and Human Behaviour, 32,* 3-5.

Schooler, J. & Engstler-Schooler, T. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*(1), 36-71.

Sears, D. O. (1986). College sophomores in the laboratory: Infl uences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51,* 515–530.

Semmler, C., Brewer, N., & Douglass, A. B. (in press). Jurors believe eyewitnesses. In B. L. Cutler (Ed.), *Conviction of the innocent: Lessons from psychological research.* Washington, DC: APA Books.

Sigall, H. & Ostrove, N. (1975) Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology, 31*(3), 410-414.

Slater, A. (1995) Saving money in ID parades. *Policing, 11*, 203-231.

Smith, S. M., Lindsay, R. C. L., Pryke, S. & Dysart, J. E. (2001) Postdictors of eyewitness errors: Can false identifications be diagnosed in the cross-race situation? *Psychology, Public Policy, and Law, 7*(1), 153-169.

Stasser, G., Kerr, N. L., & Bray, R. M. (1982). The social psychology of jury deliberations: Structure, process, and product. In N. Kerr & R. Bray (Eds.), *The psychology of the courtroom* (pp. 221-256). New York: Academic Press.

Steblay, N. (1992) A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16*(4), 413-424.

Steblay, N. (1997) Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior*, *21,* 283-297.

Steblay, N. (2006) Observations on the Illinois Lineup Data. Retrieved 12/12/2010 from http://www.psychology.iastate.edu/~glwells/Steblay_Evanston.pdf

Steblay, N. (2011) What We Know Now: The Evanston Illinois Field Lineups. *Law & Human Behaviour, 35*(1), 1-12.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001) Eyewitness Accuracy rates in Sequential and Simultaneous Lineup Presentations: A Meta-Analytic Comparison. *Law and Human Behavior, 25*(5), 459-473.

Steblay, N., Dysart, J. E. & Wells, G. L. (2011) Seventy-two tests of the Sequential Lineup Superiority Effect: A Meta-Analysis and Policy Discussion. *Psychology, Public Policy, & Law, 17*(1), 99-139.

Sporer, S. L. (2001) The cross-race effect: Beyond recognition of faces in the laboratory. *Psychology, Public Policy, and Law, 7*(1), 170-200.

Swim, J., Borgida, E., & McCoy, K. (1993). Videotaped vs. in-court witness testimony: Is protecting the child witness jeopardizing due process? *Journal of Applied Social Psychology, 23,* (8), 603-631.

Tarling, R., Dowds, L. & Budd, T. (2000) Victim and Witness Intimidation: Key Findings from the British Crime Survey. Great Britain Home Office Research Development and Statistics Directorate. Retrieved on 22/02/2007 from http://www.ncjrs.gov/App/publications/abstract.aspx?ID=189375

Tausz, D. & Ellison, L. E. (2005) Youth Court: Whether legislative provision requiring Special Measures direction to be given in relation to child witness in need of special prosecution in manner compatible with convention requirement for a fair trial. *Criminal Law Review,* June, 497-500.

Tetterton, V.S. & Warren, A.R. (2005) Using witness confidence can impair the ability to detect deception. *Criminal Justice and Behavior, 32,* 433 – 451

Tinsley, Y. (2001) Juror Decision-Making: A Look Inside the Jury Room. *The British Criminology Conference: Selected Proceedings. Volume 4.*

The Police and Criminal Evidence Act (1984) Code D (2008). Home Office. Retrieved 5 February 2009 from http://police.homeoffice.gov.uk/publications/operational-policing/2008_PACE_Code_D_(final).pdf

The Vulnerable Witnesses (Scotland) Act, 2004. Retrieved on 09/08/2006 from http://www.opsi.gov.uk/legislation/scotland/acts2004/en/aspen_20040003_en_1

The Youth Justice and Criminal Evidence Act (1999) Retrieved on 10/08/2006 from http://www.nationalarchives.gov.uk/ERORecords/HO/421/2/yjceact/yjceact.htm

Tredoux, C. & Chiroro, P. (2006) Eyewitness Testimony. In C. Tredoux, D. Foster & A. Allan. Psychology and Law. Cape Town: Juta & Company.

Tulving, E., & Thomson, D.M., (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373.

Turtle, J. W. & Wells, G. L. (1988) Children versus adults as eyewitnesses: Whose testimony holds up under cross examination? In M. W. Gruneberg et al. (Eds.), Practical aspects of memory (pp. 27-33). New York: Wiley.

Valentine, T. (2006) Forensic Facial Identification. In A. Heaton-Armstrong, E. Shepherd, G, Gudjonsson & D. Wolchover (Eds). Witness Testimony: Psychological, Investigative and Evidential Perspectives. OUP Oxford.

Valentine, T., Darling, S. & Memon, A. (2007) Do strict rules and moving images increase the reliability of sequential identification procedures? *Applied Cognitive Psychology, 21*, 933-949.

Valentine, T., Harris, N. Colom Piera, A. & Darling, S. (2003). Are police video identifications fair to African-Caribbean suspects? *Applied Cognitive Psychology, 17, 459-476.*

Valentine, T., & Heaton, P. (1999) An evaluation of the fairness of police lineups and video identifications. *Applied Cognitive Psychology, 13,* S59-S72.

Valentine, T. & Mesout, J. (2009) Eyewitness identification under stress in the London Dungeon. *Applied Cognitive Psychology*, 23, 151-161.

Valentine, T., Pickering, A. & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology, 17, 969-993.*

Vital Voices: Helping Vulnerable Witnesses Give Evidence. (2003) Retrieved on 10/08/2006 from http://www.scotland.gov.uk/Publications/2003/02/16453/18641

Walker, L., Thibault, J. & Andreoli, V. (1972) Order of Presentation at Trial. *The Yale Law Journal*, *82*(2), 216-226.

Wells, G. L. (1978) Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*, 1546-1557.

Wells, G. L. (1984) The Psychology of Lineup Identifications. *Journal of Applied Social Psychology, 14*(2), 89-101.

Wells, G. L. (1993) What do we know about eyewitness identification? *American Psychologist, 48*, 553-571.

Wells, G. L. (2001) Police Lineups: Data, Theory and Policy. *Psychology, Public Policy, and Law, 7*(4), 791-801.

Wells, G. L. (2006) Comments on the Mecklenburg Report. Retrieved 18/11/2010 from http://www.psychology.iastate.edu/~glwells/Illinois_Project_Wells_comments.pdf

Wells, G. L. (2008) Does the Sequential Lineup Reduce Accurate Identifications in Addition to Reducing Mistaken Identifications? Retrieved 14 November 2008 from: http://www.psychology.iastate.edu/~glwells/SequentialNotesonlossofhits.htm

Wells, G. L. & Bradfield, A. L. (1998) "Good, you identified the suspect:" Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83*, 360-376.

Wells, G. L., Ferguson, T. J., & Lindsay, R. C. L. (1981) The tractability of eyewitness confidence and its implication for triers of fact. *Journal of Applied Psychology, 66*, 688-696.

Wells, G. L. & Olson, E. (2003) Eyewitness identification. *Annual Review of Psychology, 54*, 277-295.

Wells, G. L., Olson, E., & Charman, S. (2002). Eyewitness identification confidence. *Current Directions in Psychological Science, 11*, 151-154.

Wells, G. L. & Seelau, E. P. (1995) Eyewitness identification: Psychological research and legal policy on lineups. *Psychology, Public Policy, and Law, 1*(4), 765-791.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero S. M., & Brimacombe, C. A. E. (1998) Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22,* 603-647.

West Yorkshire Police, V.I.P.E.R. Video Database. Retrieved on 01/03/2007 from http://www.westyorkshire.police.uk/section-item.asp?sid=6&iid=3799

Wheatcroft, J., Wagstaff, G.F., & Kebbell, M.R. (2004) The influence of courtroom questioning style on actual and perceived eyewitness confidence and accuracy. *Legal and Criminological Psychology, 9*, 83-101.

Wright, D. B., Boyd, C. E. & Tredoux, C. G. (2003) Inter-racial Contact and the Own-race Bias for Face Recognition in South Africa and England. *Applied Cognitive Psychology, 17*, 365-373.

Yarmey, A. D. & Jones, H. C. L. (1983) Is the psychology of eyewitness identification a matter of common sense? In S. Lloyd-Bostock & S. R. Clifford (Eds.), Evaluating Witness Evidence, (pp13-40). Chichester, England: Wiley.

Zander, M. and Henderson, P. (1993) *Crown Court Study* Royal Commission on Criminal Justice Research Study 19. London: HMSO.

Appendices

1. **Response sheet provided for participants in V.I.P.E.R. experiments.**
2. **Standard Police Instructions for traditional and video identity parades.**
3. **Original Trial Transcript.**
4. **Manipulated Trial Transcript.**
5. **Focus Group Response Sheet (Confidence).**
6. **Questionnaire for Pilot Study & Experiments 1-2.**
7. **Standard jury instructions regarding appointing a foreperson and verdict decisions.**

Appendix 1: Response Sheet for Identity Parade Experiments

When you have viewed the identity parade before you please indicate your decision below:

The suspect *is not* present in the identity parade ☐

I believe the suspect to be parade member number ☐

Thank you for your participation.

Please contact the experimenter for further instructions.

Appendix 2: Standard Police Instructions for Traditional and Video Identity Parades

The following exemplar statement was obtained from the Crown Office and Procurator Fiscal Office:

'The person you referred to in your statement to the police, who on 24th April 2003 in Canongate, Edinburgh exposed his penis in your presence may or may not appear in the images shown.'

This statement was altered for the purposes of the current investigations:

'The person you referred to in your statement to the police, who you witnessed stealing a wallet, may or may not appear in the images shown. You must view the entire parade twice *before you may make your decision known.'*

Appendix 3: Example of Original Trial Transcript

PROOF FOR THE PROSECUTION

Graeme Allister

Prosecutor: Would you please confirm your name and address for the court please?

GA: Mr Graeme Allister. I live at 150 Hollows Avenue, Paisley.
Prosecutor: Thank you. And could you explain your involvement with this case?
GA: I was attacked by a man in the street.
Prosecutor: Do you know that man's name?
GA: Robert Wallace.
Prosecutor: How is it you know Mr Wallace?
GA: I don't know him. I've just seen him hanging about with some other people.
Prosecutor: Thank you. Can you tell the court what happened to you on the day of May Fifteenth last year?
GA: I was going to the shops and I saw Robert Wallace. As I was going past him, he said something to me.
Prosecutor: What did he say to you?
GA: I can't remember.
Prosecutor: And what happened next?
GA: I turned around and asked him what he'd said.
Prosecutor: And then what happened?
GA: He asked me what I was looking at.
Prosecutor: Can you describe to the court what happened next?
GA: I was going into the shop when I felt him grab my collar.
Prosecutor: Mr Wallace grabbed your coat collar?
GA: Yes.
Prosecutor: With one hand or both hands?
GA: With one hand at first. Then he pulled me round because then he had a hold on my collar with both hands.
Prosecutor: Did you try to free yourself from Mr Wallace's grasp?
GA: I tried to push him arms away.
Prosecutor: Did you at any time grab Mr Wallace by the throat?
GA: No.
Prosecutor: You didn't attempt to strangle or 'choke' Mr Wallace?
GA: No.
Prosecutor: Thank you. What did Mr Wallace do next?
GA: He punched me in the face over and over.
Prosecutor: How many times do you think he punched you?
GA: I don't remember.
Prosecutor: Would you say it was more or less than five times Mr Allister?
GA: More.
Prosecutor: Thank you Mr Allister, did Mr Wallace do anything else to you?
GA: He kicked me a few times.
Prosecutor: Where did he kick you?
GA: My legs and stomach.
Prosecutor: And what happened next?
GA: I remember being on the ground.
Prosecutor: Where was Mr Wallace when you were on the ground?
GA: He was standing over me and then he ran off shouting.
Prosecutor: Thank you Mr Allister. What injuries did you sustain as a result of Mr Wallace's attack?
GA: I had two gashes on my face, a cut on my leg and a big bruise on my lower stomach. I had to go to the hospital and I got twenty-two stitches in my face.

Prosecutor: Thank you Mr Allister.

CROSS EXAMINED BY THE DEFENSE

Defence: Mr Davies, you don't appear to be too sure about your memory of this alleged attack?

GA: I was attacked for no reason. It happened so quickly.

Defence: I understand why you would claim that but how can you expect the ladies and gentlemen of the jury to believe your claims when you yourself appear to be so uncertain of your own recollections?

GA: I'm not uncertain, I don't like talking about it.

Defence: Because you are lying about what happened?

GA: I'm not lying.

Defence: So you maintain that you didn't try to strangle Mr Wallace?

GA: I did not.

Defence: Mr Allister, this is a criminal case which could affect my client for several years. Can you be absolutely certain that it was Mr Wallace who allegedly threatened you?

GA: I am fairly certain it was him.

Defence: Again Mr Allister, how can you expect the jury to make a judgement upon this case when they are presented with an uncertain account of assault?

GA: It happened so quickly It is difficult to be absolutely certain.

CROSS EXAMINATION ENDS

Appendix 4: Trial Transcript used in Mock Trial Videos

PROOF FOR THE PROSECUTION

Graeme Allister

Prosecutor: Would you please confirm your name and address for the court please?

GA: Mr Graeme Allister. I live at 150 Hollows Avenue, Paisley.
Prosecutor: Thank you. And could you explain your involvement with this case?
GA: I was attacked by a man in the street.
Prosecutor: Do you know that man's name?
GA: I think its Robert Wallace.
Prosecutor: How is it you know Mr Wallace?
GA: I…I don't really know him. I've just seen him hanging about with some other people.
Prosecutor: Thank you. Can you tell the court what happened to you on the day of May Fifteenth last year?
GA: I was going to the shops and I saw Robert Wallace.
Prosecutor: And what happened then?
GA: Well, as I was going past him, I think he said something to me.
Prosecutor: What did he say to you?
GA: I don't really remember.
Prosecutor: And what happened next?
GA: I guess I turned around and asked him what he'd said.
Prosecutor: And then what happened?
GA: It's difficult to remember, it was a while ago.
Prosecutor: I understand Mr Allister but if you could take a moment and then please tell us what happened after you turned to face Mr Wallace.
GA: (Sighs) Af…After I turned around to look at him, he asked me what I was looking at.
Prosecutor: And what did you say?
GA: I didn't say anything.
Prosecutor: Ok thank you. Can you describe to the court what happened next?
GA: I went to go into the shop when I felt him grab my collar.
Prosecutor: Mr Wallace grabbed your coat collar?
GA: Yes.
Prosecutor: With one hand or both hands?
GA: With one hand at first. Then I guess he must have pulled me round 'cause then he had a hold on my collar with both hands.
Prosecutor: What happened next Mr Allister?
GA: Well, I was scared he was going to beat me up.
Prosecutor: Did you try to free yourself from Mr Wallace's grasp?
GA: I did yes. I tried to push him arms away.
Prosecutor: Did you at any time grab Mr Wallace by the throat?
GA: No I didn't, I just wanted to get away.
Prosecutor: You didn't attempt to strangle or 'choke' Mr Wallace?
GA: No… I just wanted him to leave me alone.
Prosecutor: Thank you. What did Mr Wallace do next?
GA: He just started hitting me, punching me in the face, over and over.
Prosecutor: How many times do you think he punched you?
GA: I…I really can't remember.
Prosecutor: Would you say it was more or less than five times Mr Allister?
GA: I…I guess I would say more.
Prosecutor: Thank you Mr Allister, I understand this must be difficult for you. Did Mr Wallace do anything else to you?
GA: He kicked me a few times as well.
Prosecutor: Where did he kick you?

GA: My legs and stomach.
Prosecutor: And what happened next?
GA: He must have let me go because I remember being on the ground.
Prosecutor: Where was Mr Wallace when you were on the ground?
GA: Err…I think he was standing over me and then I think I heard him shout something at me as he was running away.
Prosecutor: Do you remember what he shouted at you?
GA: Not really. It sounded like he was threatening to come back and attack me again.
Prosecutor: Thank you Mr Allister. What injuries did you sustain as a result of Mr Wallace's attack?
GA: I had two gashes on my face, a cut on my leg and a big bruise on my lower stomach. I had to go to the hospital and I got twenty-two stitches in my face.
Prosecutor: Thank you for your patience Mr Allister.

CROSS EXAMINED BY THE DEFENSE
Defence: Mr Davies, you don't appear to be too sure about your memory of this alleged attack?

GA: I was attacked for no reason. It happened so quickly it's…it's difficult to remember.
Defence: I understand why you would claim that but how can you expect the ladies and gentlemen of the jury to believe your claims when you yourself appear to be so uncertain of your own recollections?
GA: It…It's not that I'm uncertain; I just don't like talking about it.
Defence: Because you are lying about what happened?
GA: (vehement) I'm not lying.
Defence: So you maintain that you didn't try to strangle Mr Wallace?
GA: I…I didn't!
Defence: And your claims that Mr Wallace threatened to 'come back and attack you again', how sure are your memories of that?
GA: Pretty sure, I guess.
Defence: 'Pretty sure, you guess'? Mr Allister, this is a criminal case which could affect my client for several years. Can you be absolutely certain that it was Mr Wallace who allegedly threatened you?
GA: I…I'm pretty certain it was him.
Defence: Again Mr Allister, how can you expect the jury to make a judgement upon this case when they are presented with an uncertain account of assault?
GA: I guess it happened so quickly it's difficult for me to remember exactly what happened.

CROSS EXAMINATION ENDS

Appendix 5: Focus Groups Response Sheet

Please rate how confident you perceive each of the eight witnesses you are about to see on the scales below.

1 = Very Nervous, 10 = Very Confident.

Witness 1:

_____

1 2 3 4 5 6 7 8 9 10

Witness 2:

_____

1 2 3 4 5 6 7 8 9 10

Witness 3:

_____

1 2 3 4 5 6 7 8 9 10

Witness 4:

_____

1 2 3 4 5 6 7 8 9 10

Witness 5:

_____

1 2 3 4 5 6 7 8 9 10

Witness 6:

_____

1 2 3 4 5 6 7 8 9 10

Witness 7:

_____

1 2 3 4 5 6 7 8 9 10

Witness 8:

_____

1 2 3 4 5 6 7 8 9 10

Appendix 6: Questionnaire provided to Pilot Study Participants

Please rate the witness you have just seen and heard on the scales below:

How *Credible* did you feel the witness to be? 1=Not credible at all, 10=Highly credible.

_____

1 2 3 4 5 6 7 8 9 10

How *believable* did you feel the witness to be? 1=Not believable at all, 10=Highly believable.

_____

1 2 3 4 5 6 7 8 9 10

How *confident* did the witness appear to you? 1=Very nervous, 10=Very confident.

_____

1 2 3 4 5 6 7 8 9 10

Please deliver a verdict upon the defendant based solely upon what you have seen and heard in the trial.

I find the defendant:

Guilty **Not Guilty**

Thank you for your participation.

Please contact the experimenter for further instructions.

Appendix 7: Standard instructions for jurors regarding appointing a foreperson and verdict decisions

‘Your first task as a jury is to appoint a foreperson who will be responsible for chairing deliberations and reporting your verdict to the sheriff.

During the course of your deliberations you may frequently wish to monitor the pattern of verdict responses by holding a vote. A unanimous verdict may be possible but if it becomes apparent that this is not possible then please contact the court officer (experimenter) who will inform you of the majority verdict criteria.’

‘The minimum majority will require ten jurors to be unanimously in favour of the verdict to be delivered.’